



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 115178

TO: Jeffrey Parkin
Location: REM/3D39/3C18
Art Unit: 1648
Thursday, February 26, 2004

Case Serial Number: 09/877606

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Parkin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

This Page Blank (uspto)

STIC-Biotech/ChemLib

115178

From: Parkin, Jeffrey
Sent: Tuesday, February 24, 2004 6:34 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 09/877,606

Please search **SEQ ID NOS.: 3-5, 7, and 8** v. all relevant databases, including interference. Place results on both paper and e-mail. Thanks!

JSP
AU 1648
REM 3D39
2-0908

RECEIVED
FEB 25 2004
STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/25/04
Date Completed: 3/26/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 0320
WWW/Internet: _____
Other (specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:09 ; Search time 76.6835 Seconds
(without alignments)
191.599 Million cell updates/sec

Title: US-09-877-606-3
Perfect score: 292
Sequence: 1 NHTTWLEWREINNYTSLIH.....NEQELLELKWASLWFWFI 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 290 | 99.3 | 139 | 6 | ABB84625 HIV-1 str |
| 2 | 290 | 99.3 | 139 | 6 | ABB84626 HIV-1 str |
| 3 | 290 | 99.3 | 185 | 6 | ABB84628 gp41 ecto |
| 4 | 290 | 99.3 | 268 | 2 | AAY22820 SEQ ID NO |
| 5 | 290 | 99.3 | 268 | 5 | ABG68291 Envelope |
| 6 | 290 | 99.3 | 268 | 6 | ABU57698 Human imm |
| 7 | 290 | 99.3 | 344 | 5 | ABB83400 HIV gp41 |
| 8 | 290 | 99.3 | 345 | 3 | ADC14130 HIV gp41 |
| 9 | 290 | 99.3 | 346 | 7 | ABB80284 511-856 o |
| 10 | 290 | 99.3 | 359 | 6 | ABR57594 HIV gp41 |
| 11 | 290 | 99.3 | 420 | 2 | AAR33785 Translati |
| 12 | 290 | 99.3 | 519 | 2 | AAR00181 HIV-1 env |
| 13 | 290 | 99.3 | 521 | 2 | AAR69997 HIV-1 env |
| 14 | 290 | 99.3 | 853 | 2 | AAR43066 HIV-1 gp1 |
| 15 | 290 | 99.3 | 856 | 1 | AAP60131 Sequence |
| 16 | 290 | 99.3 | 856 | 2 | AAR41025 Selective |
| 17 | 290 | 99.3 | 856 | 2 | AAR41032 Selective |
| 18 | 290 | 99.3 | 856 | 2 | AAR41031 Selective |
| 19 | 290 | 99.3 | 856 | 2 | AAR41026 Selective |
| 20 | 290 | 99.3 | 856 | 2 | AAR41028 Selective |
| 21 | 290 | 99.3 | 856 | 2 | AAR41029 Selective |
| 22 | 290 | 99.3 | 856 | 2 | AAR41030 Selective |
| 23 | 290 | 99.3 | 856 | 2 | AAR41027 Selective |
| 24 | 290 | 99.3 | 856 | 3 | AAY97072 Wild type |
| 25 | 290 | 99.3 | 856 | 4 | AAB45697 HIV-1/III |

| | | | | | |
|----|-----|------|-----|---|--------------------|
| 26 | 290 | 99.3 | 856 | 6 | ABUG3327 Human lym |
| 27 | 290 | 99.3 | 856 | 7 | ABB80283 gp41 of 1 |
| 28 | 290 | 99.3 | 856 | 8 | ADE84721 Human imm |
| 29 | 277 | 94.9 | 56 | 3 | AB114532 HIV-1 iso |
| 30 | 277 | 94.9 | 56 | 4 | AAG53860 Amino aci |
| 31 | 277 | 94.9 | 56 | 4 | AU70191 HIV viral |
| 32 | 277 | 94.9 | 138 | 3 | AAY69792 HIV-1 gp4 |
| 33 | 277 | 94.9 | 147 | 6 | ABP71557 HIV-1 gp4 |
| 34 | 277 | 94.9 | 150 | 1 | AAP80745 Sequence |
| 35 | 277 | 94.9 | 177 | 5 | ABB83401 HIV gp41 |
| 36 | 277 | 94.9 | 177 | 5 | ABB83413 HIV gp41 |
| 37 | 277 | 94.9 | 192 | 2 | AAY24098 Recombina |
| 38 | 277 | 94.9 | 200 | 7 | ADC24667 HIV gp41 |
| 39 | 277 | 94.9 | 233 | 2 | AAR93199 Recombina |
| 40 | 277 | 94.9 | 268 | 2 | AAY22822 SEQ ID NO |
| 41 | 277 | 94.9 | 268 | 2 | AAY22821 SEQ ID NO |
| 42 | 277 | 94.9 | 268 | 5 | ABG68292 Envelope |
| 43 | 277 | 94.9 | 288 | 5 | ABG68293 Envelope |
| 44 | 277 | 94.9 | 268 | 6 | ABU57699 Human imm |
| 45 | 277 | 94.9 | 268 | 6 | ABU57700 Human imm |

ALIGNMENTS

RESULT 1
ABB84625
ID ABB84625 standard; protein; 139 AA.
AC ABB84625;
XX
XX
DT 23-OCT-2003 (revised)
DT 04-FEB-2003 (first entry)
XX
DE HIV-1 strain HXB2 HIVIIB gp41 ectodomain fragment.
XX
KW Envelope glycoprotein; gp41; fusion protein; haemagglutinin protein;
KW immunogen; vaccine; IgG; IgA; immunisation.
XX
OS Human immunodeficiency virus 1.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 59..65
FT FT /note= "The Cys residues which form the disulphide bond
FT FT are substituted by Ser residues in the variant protein
FT FT represented in ABB84626"
FT Misc-difference 125
FT /note= "This residue is represented as Asp in the
FT FT specification but is represented as Asp in the
FT FT corresponding fusion protein represented in ABB84628"

WO200281655-A2.

17-OCT-2002.

27-MAR-2002; 2002WO-US009353.

28-MAR-2001; 2001US-0279383P.

(CHIL-) CHILDRENS MEDICAL CENT.

Weissenborn W, Wiley D, Mantia N, Kozlowski P;

WPI; 2003-058532/05.

Novel fusion protein for inducing human immunodeficiency virus-antigen specific IgG and IgA antibodies, has ectodomain of HIV-1 envelope glycoprotein gp41 fused to fragment of influenza virus hemagglutinin protein.

Disclosure; Page 27; 78pp; English.

This invention describes a novel fusion protein construct comprising a

CC fragment of the human immunodeficiency virus envelope glycoprotein gp41
 CC ectodomain and a second amino acid region composed of a fragment of the
 CC influenza virus haemagglutinin protein. The composition of the invention
 CC is soluble at physiological pH and is useful as an immunogen, useful in a
 CC vaccine for the induction of human immunodeficiency virus (HIV)-antigen
 CC specific serum IgG and secretory IgA antibodies in vivo. The immunogen is
 CC administered to the body of a patient as a primary immunisation, then
 CC mucosally administered as a secondary immunisation. This sequence
 CC represents the HIV-1 strain HXB2 native envelope glycoprotein gp41
 CC corresponding to residues 29-167 of the ectodomain. This native protein
 CC is modified to eliminate a disulphide bond and many residues are
 CC deglycosylated prior to the construction of the fusion protein
 CC represented in ABB84628. (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 139 AA;

Query Match 99.3%; Score 290; DB 6; Length 139;
 Best Local Similarity 98.1%; Pred. No. 1.4e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQSQOEKNEQELLELDKWSLWNWFI 52
 Db 85 NHTTWLEWDREINNTSLIHSLSIESQSQOEKNEQELLELDKWSLWNWFI 136

RESULT 2
 ABB84626
 ID ABB84626 standard; protein; 139 AA.
 AC ABB84626;
 XX
 XX
 DT 04-FEB-2003 (first entry)
 DE HIV-1 strain HXB2 mutant gp41 ectodomain fragment.
 XX
 KW Envelope glycoprotein; gp41; fusion protein; haemagglutinin protein;
 KW immunogen; vaccine; IgG; IgA; immunisation; mutant; muten.
 XX
 OS Human immunodeficiency virus 1.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 59 /label= C59S
 FT /note= "Wild type Cys is replaced by Ser"
 FT Misc-difference 65 /label= C65S
 FT /note= "Wild type Cys is replaced by Ser"
 FT Misc-difference 125 /note= "This residue is represented as Asp in the
 FT specification but is described as Asp the corresponding
 FT fusion protein represented in ABB84628"
 XX
 PN WO200281655-A2.
 XX
 PD 17-OCT-2002.
 XX
 XX 27-MAR-2002; 2002WO-US009353.
 XX
 XX 28-MAR-2001; 2001US-0279383P.
 XX
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 PA
 XX Weissenhorn W, Wiley D, Mantis N, Neutra MR, Kozlowski P;
 PI WPI; 2003-058532/05.
 XX
 DR Novel fusion protein for inducing human immunodeficiency virus-antigen
 PT specific IgG and IgA antibodies, has ectodomain of HIV-1 envelope
 PT glycoprotein gp41 fused to fragment of influenza virus haemagglutinin
 PT protein.
 XX
 PS Disclosure; Page; 78pp; English.

XX
 CC This invention describes a novel fusion protein construct comprising a
 CC fragment of the human immunodeficiency virus envelope glycoprotein gp41
 CC ectodomain and a second amino acid region composed of a fragment of the
 CC influenza virus haemagglutinin protein. The composition of the invention
 CC is soluble at physiological pH and is useful as an immunogen, useful in a
 CC vaccine for the induction of human immunodeficiency virus (HIV)-antigen
 CC specific serum IgG and secretory IgA antibodies in vivo. The immunogen is
 CC administered to the body of a patient as a primary immunisation, then
 CC mucosally administered as a secondary immunisation. This sequence
 CC represents a HIV-1 strain HXB2 mutant envelope glycoprotein gp41
 CC corresponding to residues 29-167 of the ectodomain, which is used in the
 CC construction of the fusion protein represented in ABB84628. The wild-type
 CC protein has been modified to eliminate a disulphide bond and many
 CC residues have been deglycosylated to generate the mutant. NOTE: This
 CC sequence does not appear in the disclosure and has been created using the
 CC wild-type protein represented in ABB84625 and from information given on
 CC Page 26 of the Disclosure
 XX
 SQ Sequence 139 AA;

Query Match 99.3%; Score 290; DB 6; Length 139;
 Best Local Similarity 98.1%; Pred. No. 1.4e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQSQOEKNEQELLELDKWSLWNWFI 52
 Db 85 NHTTWLEWDREINNTSLIHSLSIESQSQOEKNEQELLELDKWSLWNWFI 136

RESULT 3
 ABB84628
 ID ABB84628 standard; protein; 185 AA.
 AC ABB84628;
 XX
 XX
 DT 04-FEB-2003 (first entry)
 DE gp41 ectodomain/haemagglutinin subunit construct.
 XX
 KW Envelope glycoprotein; gp41; fusion protein; haemagglutinin protein;
 KW immunogen; vaccine; IgG; IgA; immunisation; mutant.
 XX
 OS Human immunodeficiency virus 1.
 OS Influenza virus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Region 1..137
 FT /note= "HIV-1 gp41 region"
 FT Misc-difference 59 /label= C59S
 FT /note= "gp41 wild type Cys is replaced by Ser"
 FT Misc-difference 65 /label= C65S
 FT /note= "gp41 wild type Cys is replaced by Ser"
 FT Region 141..185
 FT /note= "haemagglutinin subunit 2 region"
 FT Misc-difference 145 /note= "Encoded by ATCATC"
 XX
 PN WO200281655-A2.
 XX
 PD 17-OCT-2002.
 XX
 XX 27-MAR-2002; 2002WO-US009353.
 XX
 XX 28-MAR-2001; 2001US-0279383P.
 XX
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 PA
 XX Weissenhorn W, Wiley D, Mantis N, Neutra MR, Kozlowski P;
 PI

DR WPI; 2003-058532/05.
 DR N-PSDB; ABS56799.
 XX Novel fusion protein for inducing human immunodeficiency virus-antigen
 PT specific IgG and IgA antibodies, has ectodomain of HIV-1 envelope
 PT glycoprotein gp41 fused to fragment of influenza virus haemagglutinin
 PT protein.
 XX Disclosure; Page 30; 78pp; English.
 XX This invention describes a novel fusion protein construct comprising a
 CC fragment of the human immunodeficiency virus envelope glycoprotein gp41
 CC ectodomain and a second amino acid region composed of a fragment of the
 CC influenza virus haemagglutinin protein. The composition of the invention
 CC is soluble at physiological pH and is useful as an immunogen, useful in a
 CC vaccine for the induction of human immunodeficiency virus (HIV)-antigen
 CC specific serum IgG and secretory IgA antibodies in vivo. The immunogen is
 CC administered to the body of a patient as a primary immunisation, then
 CC mucosally administered as a secondary immunisation. This sequence
 CC represents an immunogenic fusion protein composed of a human HIV-1 strain
 CC HXB2 mutant envelope glycoprotein gp41 ectodomain fragment (ABB84626) and
 CC an influenza virus haemagglutinin protein subunit 2 fragment
 CC corresponding to residues 43-88 of the full length protein (ABB84627)
 XX
 XX Sequence 185 AA;
 SQ
 Query Match 99.3%; Score 290; DB 6; Length 185;
 Best Local Similarity 98.1%; Pred. No. 1.9e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWASLWNWFI 52
 DB 86 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWASLWNWFI 137
 RESULT 4
 AAY22820
 ID AAY22820 standard; protein; 268 AA.
 XX
 AC AAY22820;
 DT 19-AUG-1999 (first entry)
 XX
 DE SEQ ID NO. 16 from WO9820036.
 XX
 KW HIV; gp41 protein; constrained helical peptide; HIV infection; vaccine;
 KW antibody; viral membrane fusion; viral infectivity;
 KW ligand affinity purification; protein A replacement;
 KW immunoglobulin purification; epitope mimic.
 XX
 OS Human immunodeficiency virus.
 XX
 PN WO9820036-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 05-NOV-1997; 97WO-US020069.
 XX
 PR 06-NOV-1996; 96US-00743698.
 PR 16-JUN-1997; 97US-00876698.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Braisted A, Judice JK, McDowell RS, Phelan JC, Starovasnik MA;
 PI Wells JA;
 XX
 DR WPI; 1998-286866/25.
 XX
 XX Production of constrained helical peptide(s) by linking side chains on
 PT termini of octa-peptide - derived from human immunodeficiency virus gp41
 PT protein, useful in vaccines for treatment and prevention of infection.
 XX
 PS Claim 11; Page 156; 279pp; English.

XX Peptides AAY22805-Y22917 are derived from Human immunodeficiency virus
 CC (HIV). Specifically, AAY22810-Y22910 are derived from gp41 proteins of
 CC known HIV virus strains (AAY22810, AAY22871, AAY22880, AAY22888 and
 CC AAY22903 represent consensus sequences of various sections of the gp41
 CC protein). Sequences derived from the peptides are used to produce
 CC constrained helical peptides of the invention. The constrained helical
 CC peptide is produced by synthesising an octapeptide in which both terminal
 CC amino acids have a side-chain that includes a group able to form an amide
 CC bond, and cyclising the octapeptide by reacting the specified side-chain
 CC residues with a difunctional linker to produce two amide bonds. The
 CC constrained helical peptides are used to treat or prevent HIV infection,
 CC especially as vaccines that generate antibodies that prevent viral
 CC membrane fusion or infectivity. Vaccines may contain constrained helical
 CC peptides derived from several different strains of HIV. The antibodies
 CC are also useful for diagnosing HIV infection. Other uses for the
 CC constrained helical peptides are in affinity purification of ligands
 CC (particularly where complete binding protein is not readily available,
 CC e.g. replacements for protein A in immunoglobulin purification); as
 CC epitope mimics for antibody production; for isolation of synthetic
 CC antibody clones from phage display libraries, or as stable forms of
 CC "floppy" peptides or proteins
 XX
 XX Sequence 268 AA;
 SQ
 Query Match 99.3%; Score 290; DB 2; Length 268;
 Best Local Similarity 98.1%; Pred. No. 2.9e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWASLWNWFI 52
 DB 154 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWASLWNWFI 205
 RESULT 5
 ABG68291
 ID ABG68291 standard; protein; 268 AA.
 XX
 AC ABG68291;
 XX
 DT 29-AUG-2003 (revised)
 DT 07-OCT-2002 (first entry)
 XX
 DE Envelope protein gp41 from HIV clade B strain #10.
 XX
 KW HIV; glycoprotein; gp41; antigen; helical conformation;
 KW virus-induced membrane fusion; acquired immunodeficiency syndrome; AIDS;
 KW viral envelope protein; vaccine; virucide; anti-HIV.
 XX
 OS Human immunodeficiency virus 1; clade B.
 XX
 PN US6271198-B1.
 XX
 PD 07-AUG-2001.
 XX
 PF 05-NOV-1997; 97US-00965056.
 XX
 PR 06-NOV-1996; 96US-00743698.
 PR 16-JUN-1997; 97US-0049787P.
 PR 16-JUN-1997; 97US-00876698.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Braisted AC, Judice JK, McDowell RS, Phelan JC, Starovasnik MA;
 PI Wells JA;
 XX
 DR WPI; 2002-487624/52.
 XX
 XX New cyclic peptides from human immune deficiency virus gp41, useful for
 PT treatment or prevention of HIV infection, are constrained to have alpha-
 PT helical conformation.
 XX
 PS Disclosure; Col 151-154; 175pp; English.

XX The invention relates to cyclic peptides (A) with a constrained helical
 CC conformation, derived from gp41 (Glycoprotein 41, a viral envelope
 CC protein) of human immunodeficiency virus (HIV). The cyclic
 CC peptides have formulas given in the specification part of which are
 CC derived from a consensus sequence of gp41 derived from HIV clades A, B,
 CC C, D, E or O. The peptides are used to cause induction of a specific
 CC immune response, resulting in antibodies that prevent virus-induced
 CC membrane fusion. The peptides are used to treat subjects with, or at risk
 CC of, HIV infection, either as antifusion/anti-infection agents or,
 CC preferably where associated with a carrier, as an immunogen (including as
 CC vaccine) to raise antibodies. The antibodies may be used for diagnosis or
 CC prevention/treatment of HIV infection (i.e. acquired immunodeficiency
 CC syndrome, AIDS), e.g. prevention of mother-to-child transmission or in
 CC cases of health care accidents. The peptides can be based on specific HIV
 CC strains, e.g. breakthrough isolates of HIV that have developed during
 CC vaccine trials, so a combination of them should cover a wide range of
 CC protection. The present sequence is gp41 protein from a particular HIV
 CC clade used to derive a consensus sequence of gp41. (Updated on 29-AUG-
 CC 2003 to standardise OS field)
 XX Sequence 268 AA;
 SQ

Query Match 99.3%; Score 290; DB 5; Length 268;
 Best Local Similarity 98.1%; Pred. No. 2.9e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSIIIESQKQKNEQELLELDKWSLWNWNI 52
 ||||:|||||
 DB 154 NHTTWLEWDREINNYTSLIHSIIIESQKQKNEQELLELDKWSLWNWNI 205

RESULT 6
 ABUS7698
 ID ABUS7698 standard; protein; 268 AA.
 XX
 AC ABUS7698;
 XX
 DT 10-APR-2003 (first entry)
 XX
 DE Human immunodeficiency virus (HIV) envelope protein gp41 #10.
 XX
 KW Human immunodeficiency virus; HIV; vaccine; helical peptide compound;
 KW viral membrane fusion; haptens; immunogen; peptidomimetic; gp41;
 KW envelope protein.
 XX
 OS Human immunodeficiency virus.
 XX
 PN US2002151473-A1.
 XX
 PD 17-OCT-2002.

XX 15-MAY-2001; 2001US-00854816.
 XX
 PR 06-NOV-1996; 96US-00743698.
 PR 16-JUN-1997; 97US-0049787P.
 PR 16-JUN-1997; 97US-00876698.
 PR 05-NOV-1997; 97US-00965056.
 XX
 PA (BRAI/) BRAISTED A C.
 PA (JUDI/) JUDICE J K.
 PA (MCDG/) MCDOWELL R S.
 PA (PHEL/) PHELAN J C.
 PA (STAR/) STAROVASNIK M A.
 PA (WELL/) WELLS J A.
 XX
 XX Braisted AC, Judice JK, McDowell RS, Phelan JC, Starovasnik MA;
 PI Wells JA;
 PI
 XX WPI; 2003-182525/18.
 XX
 XX Novel constrained helical peptide compound useful for prophylactically or
 PT therapeutically treating mammal at risk for or infected with human

PT immunodeficiency virus.
 XX Disclosure; Fig 16; 180pp; English.
 XX
 CC The invention describes a constrained helical peptide compound (I)
 CC comprising a first constrained helical peptide comprising a sequence of 8
 CC amino acids (a.as) having a first and second terminal residue both
 CC flanking an internal sequence of 6 a.as, where the terminal residues have
 CC a side chain that are linked to each other forming a locking group to
 CC form a constrained helical peptide. (I) is useful for preparing
 CC antibodies that prevent viral membrane fusion, as haptens, preferably
 CC attached to a carrier, for use as an immunogen to raise antibodies that
 CC have a diagnostic use, as a vaccine for treatment of patients at risk of
 CC or infected with HIV, to create combinatorial constrained helical peptide
 CC libraries that are useful in chemical selection systems, to isolate the
 CC binding determinants from alpha-helical binding domains of known
 CC proteins, for determining whether a binding determinate in an alpha-
 CC helical binding domain of a known protein can serve as a structural model
 CC for the design of peptidomimetics, to replace intact binding proteins or
 CC protein binding domains in the affinity purification of ligands, to mimic
 CC epitopes in proteins to selectively raise polyclonal or monoclonal
 CC antibodies against such individual epitopes for isolating synthetic
 CC antibody clones with a selected binding activity from phage display
 CC combinatorial libraries, to provide conformationally stable variants of
 CC peptides or proteins which exhibit floppy or unstable alpha-helical
 CC secondary structure at one or more sites in unrestrained form under
 CC conditions of interest. This is the amino acid sequence of an HIV
 CC envelope protein gp41, fragments of which are used in the creation of
 CC locked helix peptides
 XX
 SQ Sequence 268 AA;

Query Match 99.3%; Score 290; DB 6; Length 268;
 Best Local Similarity 98.1%; Pred. No. 2.9e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSIIIESQKQKNEQELLELDKWSLWNWNI 52
 ||||:|||||
 DB 154 NHTTWLEWDREINNYTSLIHSIIIESQKQKNEQELLELDKWSLWNWNI 205

RESULT 7
 ABB83400
 ID ABB83400 standard; protein; 344 AA.
 XX
 AC ABB83400;
 XX
 DT 19-SEP-2002 (first entry)
 XX
 DE HIV gp41 LAI protein #1.
 XX
 KW HIV; gp41; LAI; virucide; anti-HIV; vaccine; HIV infection.
 XX
 OS Human immunodeficiency virus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Domain /label= Fusion_peptide
 FT 173..194
 FT /label= Transmembrane_domain
 XX
 XX WO200253587-A2.
 PN
 XX
 PD 11-JUL-2002.
 XX
 XX 04-JAN-2002; 2002WO-FR0000031.
 XX
 XX 05-JAN-2001; 2001FR-00000141.
 PR 23-JAN-2001; 2001FR-00000848.
 XX
 XX (AVET) AVENTIS PASTEUR.
 XX
 XX Brasseur R, Charloreaux B, Chevalier M, El Habib R, Krell T;

PI Sodyer R;
 XX WPI; 2002-528852/56.
 XX
 PT New mutant human immune deficiency virus gp41 polypeptide, useful in
 PT vaccines against human immune deficiency virus, mimics the intermediate
 PT state of the native polypeptide.
 XX
 XX Disclosure; Page 29; 29pp; French.
 XX
 CC The present invention relates to novel mutant HIV gp41 LAI proteins
 CC (ABB83411-ABB83420). The mutants can form a structure corresponding to,
 CC or mimicking, the intermediate state of gp41. The mutants, or their
 CC conjugate with a carrier, or vectors containing nucleic acid that encode
 CC them, are used in vaccines for treatment or prevention of infection by
 CC HIV. The present sequence is wild-type HIV gp41 LAI protein, which was
 CC used to generate the mutants of the invention
 XX
 XX Sequence 344 AA;
 SQ
 Query Match 99.3%; Score 290; DB 5; Length 344;
 Best Local Similarity 98.1%; Pred. No. 3.8e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNNWFI 52
 Db 113 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNNWFI 164
 RESULT 8
 ADC14130
 ID ADC14130 standard; protein; 345 AA.
 XX
 AC ADC14130;
 DT 18-DEC-2003 (first entry)
 XX
 XX HIV gp41 protein used to create the variants Trx-N and GST-C.
 DE
 XX HIV infection; gp41; immunoassay; Trx-N; thioredoxin; GST-C;
 KW glutathione-S-transferase; AIDS; HIV.
 XX
 XX Human immunodeficiency virus 1.
 OS
 XX US6294341-B1.
 PN
 XX 25-SEP-2001.
 PD
 XX 19-MAR-1999; 99US-00272342.
 PF
 XX 21-MAR-1998; 98KR-00009858.
 PR
 XX (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
 PA
 XX Yu YG, Kim S, Ryu J;
 PI WPI; 2000-585178/73.
 DR
 XX
 XX Detecting inhibitors of gp41 activity, useful for detecting inhibitors of
 PT human immunodeficiency virus (HIV) infection, by employing an assay
 PT detecting variant protein thioredoxin-N and the glutathione S-transferase
 PT protein interaction.
 PT
 XX Disclosure; SEQ ID NO 7; 13pp; English.
 PS
 XX
 XX This invention relates to a novel method for detecting the presence of a
 CC substance that inhibits HIV infection. Specifically, it refers to
 CC inhibiting activity of the HIV transmembrane protein gp41, which is
 CC responsible for the infection of HIV by causing fusion between the viral
 CC envelope membrane the cell cytoplasmic membrane. The present invention
 CC describes an immunoassay that can identify the presence of an interaction
 CC between the two helical domains of gp41 by determining the interaction of
 CC two variant proteins. These variant proteins being Trx-N (N-terminal

CC helical domain of gp41 connected to thioredoxin) and GST-C (C-terminal
 CC helical domain of gp41 linked to glutathione-S-transferase). As such, the
 CC immunoassay can be used to detect substances that bind to one of the two
 CC helical domains of gp41 thereby inhibiting gp41 activity and HIV
 CC infection, and in turn can be used to develop a therapeutic agent for
 CC AIDS. This polypeptide sequence is the HIV gp41 protein, used to create
 CC the chimeric Trx-N and GST-C variant proteins of the invention.
 XX
 XX Sequence 345 AA;
 SQ
 Query Match 99.3%; Score 290; DB 3; Length 345;
 Best Local Similarity 98.1%; Pred. No. 3.8e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNNWFI 52
 Db 113 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNNWFI 164
 RESULT 9
 ABB80284
 ID ABB80284 standard; protein; 346 AA.
 XX
 AC ABB80284;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 XX 511-856 of gp41 of 1987 HIV1 isolate.
 DE
 XX Envelope glycoprotein; gp41; 1987 HIV1; antigen; immunocomplex;
 KW detection.
 XX
 XX Human immunodeficiency virus type 1.
 OS
 XX WO2003073992-A2.
 PN
 XX 12-SEP-2003.
 PD
 XX 26-FEB-2003; 2003WO-US006206.
 PF
 XX 28-FEB-2002; 2002US-0360448P.
 PR
 XX 18-APR-2002; 2002US-0373448P.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Cloyd MW, Chen J;
 PI WPI; 2003-756752/71.
 DR
 XX Screening for HIV in a subject with a composition comprising a
 PT recombinant non-denatured HIV gp41 or gp160 antigen and detecting an
 PT immunocomplex between an antibody and the antigen.
 XX
 XX Claim 13; Page 96-97; 99pp; English.
 PS
 XX This sequence represents a fragment of the envelope glycoprotein, gp41,
 CC of 1987 HIV1 isolate. Antigenic fragments of the gp41 protein may be used
 CC in the method of the invention for screening for human immuno- deficiency
 CC virus (HIV) in a subject. The method comprises contacting a sample from
 CC the subject with a composition comprising a recombinant, non-denatured
 CC HIV gp41 antigen under conditions that permit formation of an
 CC immunocomplex between any antibody in the sample that can specifically
 CC bind to the gp41 antigen; and detecting whether an immunocomplex is
 CC formed between an antibody and the gp41 antigen
 XX
 XX Sequence 346 AA;
 SQ
 Query Match 99.3%; Score 290; DB 7; Length 346;
 Best Local Similarity 98.1%; Pred. No. 3.8e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNNWFI 52
 Db 113 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNNWFI 164

PD 22-AUG-1996.
 XX
 PF 13-FEB-1996; 96WO-US001895.
 XX
 PR 15-FEB-1995; 95US-00389459.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Morrow CD, Porter DC, Ansardi DC;
 XX
 XX WPI; 1996-393136/39.
 DR N-PSDB; AAT33295.
 XX
 PT Encapsulation of recombinant polio:virus nucleic acid for use in vaccines
 PT - using a polio:virus nucleic acid which lacks the P1 capsid region and
 PT an expression system which provides the region.
 XX
 XX Disclosure; Page 61-63; 108pp; English.
 XX
 CC cDNA sequences (AAT33293-95) respectively code for the gag, pol and env
 CC proteins (AAW00179-81) of HIV-1. They can be used to substitute the P1
 CC capsid gene of poliovirus in recombinant poliovirus nucleic acids
 CC (rPNAs). Such rPNAs are encapsidated by introduction into a host cell
 CC together with a vaccinia virus or plasmid vector encoding the poliovirus
 CC P1 capsid precursor protein. Encapsidated rPNAs are useful for genetic
 CC immunisation, stimulating an immune response to the HIV-1 protein.
 CC (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 519 AA;
 SQ
 Query Match 99.3%; Score 290; DB 2; Length 519;
 Best Local Similarity 98.1%; Pred. No. 6.1e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTLEWDREINNTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
 DB 420 NHTTLEWDREINNTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 471
 RESULT 13
 AAR69997
 ID AAR69997 standard; protein; 521 AA.
 XX
 AC AAR69997;
 XX
 DT 16-OCT-2003 (revised)
 DT 13-SEP-1995 (first entry)
 XX
 DE HIV-1 env protein.
 XX
 KW Encapsulation; Human immunodeficiency virus; type 1; immunogenic;
 KW poliovirus.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN CA2125344-A.
 XX
 XX 02-JAN-1995.
 XX
 PF 07-JUN-1994; 94CA-02125344.
 XX
 PR 01-JUL-1993; 93US-00087009.
 XX
 XX (UABR-) UAB RES FOUND.
 XX
 XX Morrow CD;
 PI
 XX WPI; 1995-099021/14.
 DR N-PSDB; AAQ80575.
 XX
 XX Method for encapsidating recombinant polio:virus nucleic acid - useful
 PT for providing compens. to stimulate immune response.
 XX

PS Disclosure; Page 38; 62pp; English.
 XX
 CC The sequence is that of the HIV-1 env protein. The DNA encoding such
 CC protein is used in a method (claimed) to encapsulate poliovirus CDNA to
 CC make it more immunogenic. See also AAR69995-6. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 521 AA;
 Query Match 99.3%; Score 290; DB 2; Length 521;
 Best Local Similarity 98.1%; Pred. No. 6.1e-23;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTLEWDREINNTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
 DB 420 NHTTLEWDREINNTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 471
 RESULT 14
 AAW43066
 ID AAW43066 standard; peptide; 853 AA.
 XX
 AC AAW43066;
 XX
 DT 17-OCT-2003 (revised)
 DT 11-SEP-1998 (first entry)
 XX
 DE HIV-1 gp120 protein fragment from isolate HXB2.
 XX
 KW gp120 protein; purification; fractionation; ion exchange; chromatography;
 KW binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5696238-A.
 XX
 PD 09-DEC-1997.
 XX
 PF 11-MAY-1995; 95US-00439286.
 XX
 PR 20-AUG-1991; 91US-00684963.
 PR 16-AUG-1993; 93US-00109002.
 PR 09-MAY-1994; 94US-00240073.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Scandella C, Haigwood NL;
 XX
 DR WPI; 1998-041353/04.
 XX
 PT Purification of HIV gp120 - using chromatographic methods.
 XX
 XX Disclosure; Fig 2A-W; 53pp; English.
 XX
 CC AAW43066-W43080 are fragments of the gp120 protein from different human
 CC immunodeficiency virus type 1 (HIV-1) isolates. These proteins are used
 CC in a novel method for purifying HIV gp120 so as to provide a purified
 CC gp120 glycoprotein having protein/protein binding properties
 CC substantially identical to natural viral HIV gp120. The method involves
 CC fractionating a crude gp120 preparation containing full-length,
 CC glycosylated gp120 using ion exchange chromatography so as to provide a
 CC first collection of fractions. A fraction from the first collection is
 CC selected that exhibits specific binding affinity for CD4 peptide, thereby
 CC producing a first fractionated material. The first fractionated material
 CC is fractionated by hydrophobic interaction chromatography so as to
 CC provide a second collection of fractions from which a second collection
 CC is selected that exhibits specific binding affinity for CD4 peptide. This
 CC second fraction is fractionated by size exclusion chromatography so as to
 CC provide a third collection of fractions exhibiting specific binding
 CC affinity for CD4 peptide, thereby providing the purified gp120. The
 CC purified gp120 can be used for antibody production and in vaccines.
 XX (Updated on 17-OCT-2003 to standardise OS field)
 XX

```

SQ Sequence 853 AA;
  Query Match          99.3%; Score 290; DB 2; Length 853;
  Best Local Similarity 98.1%; Pred. No. 1.1e-22;
  Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQOQEKNEQELLELDKWSLWNWFNI 52
   ||||:|||||
Db 621 NHTTWLEWDREINNTSLIHSLSIESQOQEKNEQELLELDKWSLWNWFNI 672

RESULT 15
AAP60131
ID AAP60131 standard; protein; 856 AA.
XX
AC AAP60131;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-JUN-1991 (first entry)
XX
DE Sequence of the AIDS envelope protein.
XX
KW AIDS; HIV; LAV; HTLV-III; vaccine; antibody; epitope; antigen; diagnosis.
XX
OS Human T-cell lymphotropic virus 3.
XX
PN EP199301-A.
XX
PD 29-OCT-1986.
XX
PF 18-APR-1986; 86EP-00105371.
XX
PR 19-APR-1985; 85US-00725021.
XX
PA (HOFF ) HOFFMANN-LA ROCHE AG.
PA (USGO ) US GOVERNMENT.
PA (HEAL-) DEPT HEALTH & HUMAN SERV.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wongstaal FY;
XX
WPI; 1986-286067/44.
DR N-PSDB; AAN60128.
XX
PT Envelope protein of acquired immune deficiency syndrome virus - useful
PT for improved testing of human blood for antibodies against virus and as
PT antigen for vaccines.
XX
PS Claim 33; Fig 6A; 46pp; English.
XX
CC An expression vector contg. a gene (AAN60128) coding for an envelope
CC protein of an AIDS virus, and the envelope protein of an AIDS virus
CC (AAP60131) are claimed. The vector is pref. a member of the pEV/env
CC family, e.g. pEV1.2 or 3/env 44-640 or 205-640. (Updated on 25-MAR-2003
CC to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 856 AA;
  Query Match          99.3%; Score 290; DB 1; Length 856;
  Best Local Similarity 98.1%; Pred. No. 1.1e-22;
  Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQOQEKNEQELLELDKWSLWNWFNI 52
   ||||:|||||
Db 624 NHTTWLEWDREINNTSLIHSLSIESQOQEKNEQELLELDKWSLWNWFNI 675
  
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:37:04 ; Search time 20.7342 Seconds
(without alignments)
129.475 Million cell updates/sec

Title: US-09-877-606-3
Perfect score: 292
Sequence: 1 NHTTWLEWDREINNYTSLIH.....NQELLELDKWSLWNNWFI 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/2/iaa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 290 | 99.3 | 138 | 4 | US-09-570-921-20 |
| 2 | 290 | 99.3 | 138 | 4 | US-09-570-921-21 |
| 3 | 290 | 99.3 | 268 | 3 | US-08-965-056-16 |
| 4 | 290 | 99.3 | 345 | 3 | US-09-272-342B-7 |
| 5 | 290 | 99.3 | 519 | 1 | US-08-589-446-8 |
| 6 | 290 | 99.3 | 519 | 1 | US-08-444-882-8 |
| 7 | 290 | 99.3 | 519 | 2 | US-08-389-459A-8 |
| 8 | 290 | 99.3 | 519 | 3 | US-08-967-867A-8 |
| 9 | 290 | 99.3 | 856 | 2 | US-07-916-098A-2 |
| 10 | 290 | 99.3 | 856 | 4 | US-09-337-387-11 |
| 11 | 277 | 94.9 | 56 | 4 | US-09-779-451-4 |
| 12 | 277 | 94.9 | 138 | 4 | US-09-570-921-22 |
| 13 | 277 | 94.9 | 138 | 4 | US-09-570-921-24 |
| 14 | 277 | 94.9 | 138 | 4 | US-09-570-921-26 |
| 15 | 277 | 94.9 | 138 | 4 | US-09-570-921-58 |
| 16 | 277 | 94.9 | 237 | 3 | US-08-388-353-641 |
| 17 | 277 | 94.9 | 237 | 3 | US-08-488-551B-641 |
| 18 | 277 | 94.9 | 268 | 3 | US-08-965-056-17 |
| 19 | 277 | 94.9 | 268 | 3 | US-08-965-056-18 |
| 20 | 277 | 94.9 | 282 | 5 | PCT-US95-13335-1 |
| 21 | 277 | 94.9 | 338 | 3 | US-08-486-099-90 |
| 22 | 277 | 94.9 | 338 | 3 | US-08-360-107A-100 |
| 23 | 277 | 94.9 | 338 | 3 | US-08-484-223B-90 |
| 24 | 277 | 94.9 | 338 | 3 | US-08-919-597-90 |
| 25 | 277 | 94.9 | 338 | 3 | US-08-475-668A-90 |
| 26 | 277 | 94.9 | 338 | 3 | US-08-485-551A-90 |
| 27 | 277 | 94.9 | 338 | 3 | US-08-471-913A-90 |

| | | | | | | |
|----|-----|------|-----|---|-------------------|-------------------|
| 28 | 277 | 94.9 | 338 | 3 | US-08-485-264A-90 | Sequence 90, Appl |
| 29 | 277 | 94.9 | 338 | 4 | US-08-474-349A-90 | Sequence 90, Appl |
| 30 | 277 | 94.9 | 338 | 4 | US-08-255-208A-26 | Sequence 26, Appl |
| 31 | 277 | 94.9 | 338 | 4 | US-08-470-896-90 | Sequence 90, Appl |
| 32 | 277 | 94.9 | 338 | 4 | US-08-485-546A-90 | Sequence 90, Appl |
| 33 | 277 | 94.9 | 345 | 4 | US-08-817-441-49 | Sequence 49, Appl |
| 34 | 277 | 94.9 | 345 | 4 | US-09-779-451-8 | Sequence 8, Appl |
| 35 | 277 | 94.9 | 615 | 3 | US-09-257-490-11 | Sequence 11, Appl |
| 36 | 277 | 94.9 | 826 | 1 | US-08-375-510-2 | Sequence 2, Appl |
| 37 | 277 | 94.9 | 826 | 2 | US-08-487-657-2 | Sequence 2, Appl |
| 38 | 277 | 94.9 | 839 | 3 | US-08-472-240A-10 | Sequence 10, Appl |
| 39 | 277 | 94.9 | 854 | 4 | US-09-309-572-23 | Sequence 23, Appl |
| 40 | 277 | 94.9 | 854 | 4 | US-09-718-096-23 | Sequence 23, Appl |
| 41 | 277 | 94.9 | 856 | 3 | US-09-124-900-9 | Sequence 9, Appl |
| 42 | 277 | 94.9 | 861 | 1 | US-08-127-499A-14 | Sequence 14, Appl |
| 43 | 277 | 94.9 | 861 | 1 | US-08-482-847-14 | Sequence 14, Appl |
| 44 | 277 | 94.9 | 861 | 3 | US-07-956-483-10 | Sequence 10, Appl |
| 45 | 277 | 94.9 | 861 | 3 | US-07-956-483-16 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1

US-09-570-921-20
; Sequence 20, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-20

Query Match 99.3%; Score 290; DB 4; Length 138;
Best Local Similarity 98.1%; Pred. No. 7.2e+26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLIEESQOQKNEQELLELDKWSLWNNWFI 52
|||||
DB 85 NHTTWLEWDREINNYTSLIHSLIEESQOQKNEQELLELDKWSLWNNWFI 136

RESULT 2

US-09-570-921-21
; Sequence 21, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21

```
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-21

Query Match          99.3%; Score 290; DB 4; Length 138;
Best Local Similarity 98.1%; Pred. No. 7.2e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNNFNI 52
    |||||:|||||
Db 85 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNNFNI 136

RESULT 3
US-09-965-056-16
; Sequence 16, Application US/08965056
; Patent No. 6271198
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; APPLICANT: J. Kevin Judice
; APPLICANT: Robert S. McDowell
; APPLICANT: J. Christopher Phelan
; APPLICANT: Melissa A. Starovasmik
; APPLICANT: James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; TITLE OF INVENTION: Making Same
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,056
; FILING DATE: 05-No. 6271198-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-965-056-16

Query Match          99.3%; Score 290; DB 3; Length 268;
Best Local Similarity 98.1%; Pred. No. 1.6e-25;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNNFNI 52
    |||||:|||||
Db 154 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNNFNI 205

RESULT 4
US-09-272-342B-7
; Sequence 7, Application US/09272342B
; Patent No. 6294341
; GENERAL INFORMATION:
; APPLICANT: YU, YEON-GYU
; APPLICANT: KIM, SUNG-HOU
; APPLICANT: RYU, JAE-RYEON
; TITLE OF INVENTION: METHOD FOR DETECTING A SUBSTANCE HAVING AN ACTIVITY TO
; TITLE OF INVENTION: INHIBIT HIV INFECTION USING IMMUNOASSAY AND VARIANT
; FILE REFERENCE: 2901-0125-0
; CURRENT APPLICATION NUMBER: US/09/272,342B
; CURRENT FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-272-342B-7

Query Match          99.3%; Score 290; DB 3; Length 345;
Best Local Similarity 98.1%; Pred. No. 2.1e-25;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNNFNI 52
    |||||:|||||
Db 113 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNNFNI 164

RESULT 5
US-08-589-446-8
; Sequence 8, Application US/08589446
; Patent No. 5614413
; GENERAL INFORMATION:
; APPLICANT: Morrow, Casey D.
; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,446
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,009
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Geary III, William C.
; REGISTRATION NUMBER: 31,359
; REFERENCE/DOCKET NUMBER: UAG-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 519 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-589-446-8

Query Match          99.3%; Score 290; DB 1; Length 519;
Best Local Similarity 98.1%; Pred. No. 3.3e-25;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

; TELECOMMUNICATION INFORMATION:

; TELECOMMUNICATION INFORMATION:


```
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 138
; TYPE: prt
; ORGANISM: Human
US-09-570-921-26

Query Match          94.9%;      Score 277;   DB 4;    Length 138;
Best Local Similarity 94.2%;      Pred. No. 2.2e-24;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

RESULT 15
US-09-570-921-58
; Sequence 58, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRIS, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213

```

; FILE REFERENCES: I06213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 58
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-58

Query Match          94.9%   Score 277;   DB 4;   Length 138;
Best Local Similarity 94.2%   Pred. No. 2.2e-24;
Matches 49;   Conservative 2;   Mismatches 1;   Indels 0;   Gaps 0;

```

Search completed: February 25, 2004, 15:46:54
Job time : 22.7342 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: February 25, 2004, 15:44:50 ; Search time 43.443 Seconds
(without alignments)
252.744 Million cell updates/sec

Title: US-09-877-606-3
Perfect score: 292
Sequence: 1 NHTTWLEWDREINNYTSLIH.....NEQELLELDKWSLMMWNI 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 290 | 99.3 | 268 | 9 | US-09-854-816-16 |
| 2 | 290 | 99.3 | 344 | 14 | US-10-040-349B-1 |
| 3 | 290 | 99.3 | 359 | 14 | US-10-214-670-58 |
| 4 | 290 | 99.3 | 519 | 9 | US-09-756-551A-8 |
| 5 | 290 | 99.3 | 856 | 9 | US-09-476-242-1 |
| 6 | 290 | 99.3 | 856 | 14 | US-10-196-515-11 |
| 7 | 277 | 94.9 | 56 | 9 | US-09-779-451-4 |
| 8 | 277 | 94.9 | 177 | 14 | US-10-040-349B-2 |
| 9 | 277 | 94.9 | 200 | 14 | US-10-263-103-25 |
| 10 | 277 | 94.9 | 200 | 15 | US-10-438-691-8 |
| 11 | 277 | 94.9 | 221 | 14 | US-10-059-271-84 |
| 12 | 277 | 94.9 | 232 | 14 | US-10-059-271-81 |
| 13 | 277 | 94.9 | 254 | 14 | US-10-059-271-82 |
| 14 | 277 | 94.9 | 256 | 14 | US-10-059-271-97 |
| 15 | 277 | 94.9 | 268 | 9 | US-09-854-816-17 |

| | | | | | | |
|----|-----|------|------|----|-------------------|--------------------|
| 16 | 277 | 94.9 | 268 | 9 | US-09-854-816-18 | Sequence 18, Appl |
| 17 | 277 | 94.9 | 345 | 9 | US-09-779-451-8 | Sequence 8, Appl |
| 18 | 277 | 94.9 | 345 | 14 | US-10-026-741-49 | Sequence 49, Appl |
| 19 | 277 | 94.9 | 391 | 14 | US-10-059-271-93 | Sequence 93, Appl |
| 20 | 277 | 94.9 | 853 | 13 | US-10-003-035-33 | Sequence 33, Appl |
| 21 | 277 | 94.9 | 853 | 14 | US-10-286-332A-33 | Sequence 33, Appl |
| 22 | 277 | 94.9 | 853 | 15 | US-10-280-915-33 | Sequence 33, Appl |
| 23 | 277 | 94.9 | 861 | 14 | US-10-026-741-103 | Sequence 103, Appl |
| 24 | 277 | 94.9 | 1101 | 13 | US-10-003-035-53 | Sequence 53, Appl |
| 25 | 277 | 94.9 | 1101 | 14 | US-10-286-332A-53 | Sequence 53, Appl |
| 26 | 277 | 94.9 | 1101 | 15 | US-10-280-915-53 | Sequence 53, Appl |
| 27 | 277 | 94.9 | 1186 | 13 | US-10-003-035-55 | Sequence 55, Appl |
| 28 | 277 | 94.9 | 1186 | 14 | US-10-286-332A-55 | Sequence 55, Appl |
| 29 | 277 | 94.9 | 1186 | 15 | US-10-280-915-55 | Sequence 55, Appl |
| 30 | 274 | 93.8 | 268 | 9 | US-09-854-816-19 | Sequence 19, Appl |
| 31 | 267 | 91.4 | 58 | 14 | US-10-252-136-11 | Sequence 11, Appl |
| 32 | 262 | 89.7 | 726 | 14 | US-10-196-515-3 | Sequence 3, Appl |
| 33 | 262 | 89.7 | 759 | 14 | US-10-196-515-12 | Sequence 12, Appl |
| 34 | 261 | 89.4 | 48 | 14 | US-10-351-641-547 | Sequence 547, Appl |
| 35 | 260 | 89.0 | 269 | 9 | US-09-854-816-46 | Sequence 46, Appl |
| 36 | 259 | 88.7 | 268 | 9 | US-09-854-816-13 | Sequence 13, Appl |
| 37 | 256 | 87.7 | 46 | 9 | US-09-779-451-41 | Sequence 41, Appl |
| 38 | 255 | 87.3 | 269 | 9 | US-09-854-816-43 | Sequence 43, Appl |
| 39 | 254 | 87.0 | 268 | 9 | US-09-854-816-9 | Sequence 9, Appl |
| 40 | 253 | 86.6 | 269 | 9 | US-09-854-816-12 | Sequence 12, Appl |
| 41 | 252 | 86.3 | 269 | 9 | US-09-854-816-28 | Sequence 28, Appl |
| 42 | 251 | 86.0 | 268 | 9 | US-09-854-816-26 | Sequence 26, Appl |
| 43 | 251 | 86.0 | 619 | 9 | US-09-891-609-4 | Sequence 4, Appl |
| 44 | 251 | 86.0 | 646 | 9 | US-09-891-609-2 | Sequence 2, Appl |
| 45 | 251 | 86.0 | 842 | 14 | US-10-190-435-2 | Sequence 2, Appl |

ALIGNMENTS

RESULT 1

US-09-854-816-16
; Sequence 16, Application US/09854816
; Patent No. US20020151473A1

GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovasinik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSER: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816

FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, PhD., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-854-816-16

Query Match          99.3%; Score 290; DB 9; Length 268;
Best Local Similarity 98.1%; Pred. No. 3.5e-24;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 52
Db 154 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 205

RESULT 2
US-10-040-349B-1
; Sequence 1, Application US/10040349B
; Publication No. US20030082521A1
; GENERAL INFORMATION:
; APPLICANT: Brasseur, Robert
; APPLICANT: Charlotiaux, Benoit
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; TITLE OF INVENTION: Polypeptide Inducing Antibodies Neutralizing HIV
; FILE REFERENCE: 01-078-A
; CURRENT APPLICATION NUMBER: US/10/040,349B
; CURRENT FILING DATE: 2002-07-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: (1)..(344)
; OTHER INFORMATION: gp41 LAI protein
US-10-040-349B-1

Query Match          99.3%; Score 290; DB 14; Length 344;
Best Local Similarity 98.1%; Pred. No. 4.6e-24;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 52
Db 113 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 164

RESULT 3
US-10-214-670-58
; Sequence 58, Application US/10214670
; Publication No. US20030180715A1
; GENERAL INFORMATION:
; APPLICANT: Tibotec Pharmaceuticals Ltd.
; TITLE OF INVENTION: Methods and means for assessing HIV envelope inhibitor
; FILE REFERENCE: VIP-0021 seq listing
; CURRENT APPLICATION NUMBER: US/10/214,670
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: EP 01203011.0
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/310497
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 359

```

```

; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-214-670-58

Query Match          99.3%; Score 290; DB 14; Length 359;
Best Local Similarity 98.1%; Pred. No. 4.8e-24;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 52
Db 135 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 186

RESULT 4
US-09-756-551A-8
; Sequence 8, Application US/09756551A
; Patent No. US20020051768A1
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
; FILING DATE: 08-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/376,184
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 519 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-756-551A-8

Query Match          99.3%; Score 290; DB 9; Length 519;
Best Local Similarity 98.1%; Pred. No. 7.1e-24;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKQWASLWNWFI 52

```



```
|||||:|||||
Db 420 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 471

RESULT 5
US-09-476-242-1
; Sequence 1, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-1

Query Match 99.3%; Score 290; DB 9; Length 856;
Best Local Similarity 98.1%; Pred. No. 1.2e-23;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
Db 624 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 675

RESULT 6
US-10-196-515-11
; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

Query Match 99.3%; Score 290; DB 14; Length 856;
Best Local Similarity 98.1%; Pred. No. 1.2e-23;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
Db 624 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 675

RESULT 7
US-09-779-451-4
; Sequence 4, Application US/09779451
; Patent No. US20020094521A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wild, Carl T.
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors
; FILE REFERENCE: 1900.0300003
; CURRENT APPLICATION NUMBER: US/09/779,451
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/235,901
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/181,543
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-779-451-4

Query Match 94.9%; Score 277; DB 9; Length 56;
Best Local Similarity 94.2%; Pred. No. 1.8e-23;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
Db 2 NNTWMEWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 53

RESULT 8
US-10-040-349B-2
; Sequence 2, Application US/10040349B
; Publication No. US20030082521A1
; GENERAL INFORMATION:
; APPLICANT: Brasseur, Robert
; APPLICANT: Charlotaux, Benoit
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; TITLE OF INVENTION: Polypeptide Inducing Antibodies Neutralizing HIV
; FILE REFERENCE: 01-078-A
; CURRENT APPLICATION NUMBER: US/10/040,349B
; CURRENT FILING DATE: 2002-07-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; NAME/KEY: Peptide
; LOCATION: (1)..(177)
; OTHER INFORMATION: polypeptide derived from gp41 LAI
US-10-040-349B-2

Query Match 94.9%; Score 277; DB 14; Length 177;
Best Local Similarity 94.2%; Pred. No. 6.1e-23;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NHTTWLWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
Db 90 NNTWMEWDRREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 141

RESULT 9
US-10-263-103-25
; Sequence 25, Application US/10263103
; Publication No. US20030138445A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PASTEUR
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; APPLICANT: Sodoyer, Regis
; TITLE OF INVENTION: gp41 antigen
```


Search completed: February 25, 2004, 16:04:24
Job time : 44.443 secs

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|--------|----|-----------|-------------|
| | | Match | Length | | | |
| 1 | 290 | 99.3 | 856 | 1 | ENV_HV1H2 | human |
| 2 | 290 | 99.3 | 856 | 1 | ENV_HV1H3 | human |
| 3 | 290 | 99.3 | 856 | 1 | ENV_HV1LW | human |
| 4 | 277 | 94.9 | 851 | 1 | ENV_HV1B8 | human |
| 5 | 277 | 94.9 | 856 | 1 | ENV_HV1B1 | human |
| 6 | 277 | 94.9 | 861 | 1 | ENV_HV1B3 | human |
| 7 | 274 | 93.8 | 853 | 1 | ENV_HV1M5 | human |
| 8 | 268 | 91.8 | 856 | 1 | ENV_HV1P1 | human |
| 9 | 254 | 87.0 | 856 | 1 | ENV_HV1L1 | human |
| 10 | 253 | 86.6 | 856 | 1 | ENV_HV1S3 | human |
| 11 | 252 | 86.3 | 852 | 1 | ENV_HV1S1 | human |
| 12 | 251 | 86.0 | 847 | 1 | ENV_HV1S2 | human |
| 13 | 243 | 83.2 | 847 | 1 | ENV_HV1W2 | human |
| 14 | 242 | 82.9 | 856 | 1 | ENV_HV1M1 | human |
| 15 | 241 | 82.5 | 853 | 1 | ENV_HV1Z2 | human |
| 16 | 241 | 82.5 | 855 | 1 | ENV_HV1Z6 | human |
| 17 | 241 | 82.5 | 867 | 1 | ENV_HV1J3 | human |
| 18 | 240 | 82.2 | 865 | 1 | ENV_HV1R8 | human |
| 19 | 238 | 81.5 | 852 | 1 | ENV_HV1B9 | human |
| 20 | 237 | 81.2 | 853 | 1 | ENV_HV1E2 | human |
| 21 | 236 | 80.8 | 855 | 1 | ENV_HV1A2 | human |
| 22 | 236 | 80.8 | 861 | 1 | ENV_HV1K8 | human |
| 23 | 234 | 80.1 | 846 | 1 | ENV_HV1ND | human |
| 24 | 233 | 79.8 | 868 | 1 | ENV_HV1C4 | human |
| 25 | 232 | 79.5 | 843 | 1 | ENV_HV1Y2 | human |
| 26 | 232 | 79.5 | 848 | 1 | ENV_HV1J9 | human |
| 27 | 232 | 79.5 | 855 | 1 | ENV_HV1O9 | human |
| 28 | 230 | 78.8 | 863 | 1 | ENV_HV1Z8 | human |
| 29 | 229 | 78.4 | 859 | 1 | ENV_HV1MA | human |
| 30 | 213 | 72.9 | 856 | 1 | ENV_HV1Z3 | human |
| 31 | 198 | 67.8 | 854 | 1 | ENV_SIVC2 | chimpanzee |
| 32 | 142 | 48.6 | 854 | 1 | ENV_SIVAI | simian |
| 33 | 130 | 44.5 | 882 | 1 | ENV_STVM1 | simian |

| | | | | |
|------|--|-----------------------------------|------|-------------------------------------|
| ID | ENV HV1W | STANDARD; | PRT; | 856 AA. |
| AC | Q70626; | | | |
| DT | 15-JUL-1998 | (Rel. 36, Created) | | |
| DT | 15-JUL-1998 | (Rel. 36, Last sequence update) | | |
| DT | 10-OCT-2003 | (Rel. 42, Last annotation update) | | |
| DE | Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]. | | | |
| DE | GN | | | |
| OS | Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1). | | | |
| OS | Viruses; Retroid viruses; Retroviridae; Lentivirus. | | | |
| NCBI | NCBI_TaxID=82834; | | | |
| RN | [1] | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RP | MEDLINE=95127297; PubMed=7826699; | | | |
| RX | Reitz M.C. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M., Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.; | | | |
| RT | "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HTLV type IIIB)."; | | | |
| RL | AIDS Res. Hum. Retroviruses 10:1143-1155(1994). | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| CC | EMBL; U12055; AAA76690.1; -- | | | |
| DR | PDB; 1IF3; 02-MAY-01. | | | |
| DR | GlycosuitedB; Q70626; -- | | | |
| DR | InterPro; IPR000328; Env GP41. | | | |
| DR | InterPro; IPR000777; GP120. | | | |
| DR | Pfam; PF00516; GP120; 1. | | | |
| DR | Pfam; PF00517; GP41; 1. | | | |
| KW | ADDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure. | | | |
| KW | 3D-structure. | | | |
| FT | SIGNAL | 1 | 30 | |
| FT | CHAIN | 31 | 511 | |
| FT | CHAIN | 512 | 856 | EXTERIOR MEMBRANE GLYCOPROTEIN. |
| FT | DISULFID | 54 | 74 | TRANSMEMBRANE GLYCOPROTEIN. |
| FT | DISULFID | 119 | 205 | BY SIMILARITY. |
| FT | DISULFID | 126 | 196 | BY SIMILARITY. |
| FT | DISULFID | 131 | 157 | BY SIMILARITY. |
| FT | DISULFID | 218 | 247 | BY SIMILARITY. |
| FT | DISULFID | 228 | 239 | BY SIMILARITY. |
| FT | DISULFID | 296 | 331 | BY SIMILARITY. |
| FT | DISULFID | 378 | 445 | BY SIMILARITY. |
| FT | DISULFID | 385 | 418 | BY SIMILARITY. |
| FT | CARBOHYD | 88 | 88 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 136 | 136 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 141 | 141 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 156 | 156 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 160 | 160 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 186 | 186 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 197 | 197 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 230 | 230 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 234 | 234 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 241 | 241 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 262 | 262 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 276 | 276 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 289 | 289 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 295 | 295 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 301 | 301 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 332 | 332 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 339 | 339 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 356 | 356 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 386 | 386 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 392 | 392 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 397 | 397 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 406 | 406 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 448 | 448 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |
| FT | CARBOHYD | 463 | 463 | N-LINKED (GLCNAC. .) (POTENTIAL. .) |

| | | | | | |
|--|---|---|------|---------------------------------|-------------|
| FT | CARBOHYD | 611 | 611 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| FT | CARBOHYD | 616 | 616 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| FT | CARBOHYD | 624 | 624 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| FT | CARBOHYD | 637 | 637 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| FT | CARBOHYD | 674 | 674 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| FT | CARBOHYD | 750 | 750 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| FT | CARBOHYD | 816 | 816 | N-LINKED (GLCNAC. . .) | (POTENTIAL) |
| SQ | SEQUENCE | 856 AA; 96939 MW; 0C241332C7B76687 CRC64; | | | |
| Query Match 99.3%; Score 290; DB 1; Length 856; | | | | | |
| Best Local Similarity 98.1%; Pred No. 3.4e-24; | | | | | |
| Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0 | | | | | |
| Qy | 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQBELLELDKQASLWNNFNI 675 | | | | |
| Db | 624 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQBELLELDKQASLWNNFNI 675 | | | | |
| RESULT 4 | | | | | |
| ENV | ENV_HV1B8 | | | | |
| ID | ENV_HV1B8 | STANDARD; | PRT; | 851 | AA. |
| AC | PO4582; | | | | |
| DT | 13-AUG-1987 | (Rel. 05, Created) | | | |
| DT | 13-AUG-1987 | (Rel. 05, Last sequence update) | | | |
| DT | 10-OCT-2003 | (Rel. 42, Last annotation update) | | | |
| DE | Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]. | | | | |
| DE | ENV. | | | | |
| GN | Human immunodeficiency virus type 1 (BHS isolate) (HIV-1). | | | | |
| OS | Viruses; Retroid viruses; Retroviridae; Lentivirus. | | | | |
| NCBI | TaxID=11684; | | | | |
| ON | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RP | MEDLINE=85111123; PubMed=2578615; | | | | |
| RA | Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., | | | | |
| RA | Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., | | | | |
| RA | Bohmeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L., | | | | |
| RA | Laurentberger J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C. | | | | |
| RA | Wang-Staal F. | | | | |
| RT | "Complete nucleotide sequence of the AIDS virus, HTLV-III."; | | | | |
| RT | Nature 313:277-284(1985). | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a colla- | | | | |
| CC | tion by the Swiss Institute of Bioinformatics and the EMBL out- | | | | |
| CC | the European Bioinformatics Institute. There are no restriction | | | | |
| CC | use by non-profit institutions as long as its content is in | | | | |
| CC | modified and this statement is not removed. Usage by and for c- | | | | |
| CC | entitles requires a license agreement (See http://www.isb-sib.ch/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | EMBL; K02011; AAA44661.1; -- | | | | |
| DR | PDB; 1DH; 13-JAN-99. | | | | |
| DR | PDB; 1HHG; 31-OCT-93. | | | | |
| DR | PDB; 1QO3; 02-JAN-00. | | | | |
| DR | PDB; 1S2T; 24-DEC-97. | | | | |
| DR | HIV; K02011; ENVSH8. | | | | |
| DR | GlycoSuiteDB; P04582; -- | | | | |
| DR | InterPro; IPR00328; ENV GP41. | | | | |
| DR | InterPro; IPR000777; GP120. | | | | |
| DR | Pfam; PF00516; GP120; 1. | | | | |
| DR | Pfam; PF00517; GP41; 1. | | | | |
| KW | AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Sig- | | | | |
| KW | 3D-structure. | | | | |
| FT | SIGNAL | 1 | 30 | | |
| FT | CHAIN | 31 | 506 | EXTERIOR MEMBRANE GLYCOPROTEIN. | |
| FT | CHAIN | 507 | 851 | TRANSMEMBRANE GLYCOPROTEIN. | |
| FT | DISULFID | 54 | 74 | BY SIMILARITY. | |
| FT | DISULFID | 119 | 205 | BY SIMILARITY. | |
| FT | DISULFID | 126 | 196 | BY SIMILARITY. | |
| FT | DISULFID | 131 | 157 | BY SIMILARITY. | |
| FT | DISULFID | 218 | 247 | BY SIMILARITY. | |
| FT | DISULFID | 228 | 239 | BY SIMILARITY. | |
| FT | DISULFID | 296 | 331 | BY SIMILARITY. | |

FT DISULFID 378 440 BY SIMILARITY. (POTENTIAL).
 FT DISULFID 385 413 BY SIMILARITY. (POTENTIAL).
 FT CARBOHYD 88 (POTENTIAL).
 FT CARBOHYD 136 (POTENTIAL).
 FT CARBOHYD 141 (POTENTIAL).
 FT CARBOHYD 156 (POTENTIAL).
 FT CARBOHYD 160 (POTENTIAL).
 FT CARBOHYD 186 (POTENTIAL).
 FT CARBOHYD 197 (POTENTIAL).
 FT CARBOHYD 230 (POTENTIAL).
 FT CARBOHYD 234 (POTENTIAL).
 FT CARBOHYD 241 (POTENTIAL).
 FT CARBOHYD 262 (POTENTIAL).
 FT CARBOHYD 276 (POTENTIAL).
 FT CARBOHYD 295 (POTENTIAL).
 FT CARBOHYD 301 (POTENTIAL).
 FT CARBOHYD 332 (POTENTIAL).
 FT CARBOHYD 339 (POTENTIAL).
 FT CARBOHYD 356 (POTENTIAL).
 FT CARBOHYD 386 (POTENTIAL).
 FT CARBOHYD 392 (POTENTIAL).
 FT CARBOHYD 401 (POTENTIAL).
 FT CARBOHYD 443 (POTENTIAL).
 FT CARBOHYD 458 (POTENTIAL).
 FT CARBOHYD 606 (POTENTIAL).
 FT CARBOHYD 611 (POTENTIAL).
 FT CARBOHYD 620 (POTENTIAL).
 FT CARBOHYD 632 (POTENTIAL).
 FT CARBOHYD 669 (POTENTIAL).
 FT CARBOHYD 745 (POTENTIAL).
 FT CARBOHYD 811 (POTENTIAL).
 SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 94.9%; Score 277; DB 1; Length 851;
 Best Local Similarity 94.2%; Pred. No. 9.1e-23;
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSIEESQKQKNEQELLELDKWASLWNWFI 52
 DB 619 NNMTWMEWDREINNYTSLIHSIEESQKQKNEQELLELDKWASLWNWFI 670

RESULT 5
 ENV_HV1B1 STANDARD; PRT; 856 AA.
 AC F03375;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP150 precursor [Contains: Exterior membrane glycoprotein (gp120); Transmembrane glycoprotein (gp41)].
 DE Glycoprotein (gp120);
 GN ENV.
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
 OS Viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85111123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
 RA Lautenberger J.A., Papas T., Graybe J., Chang N.T., Gallo R.C.,
 RA Wong-Staal F.;
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 RL Nature 313:277-284 (1985).
 RN [2]
 RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=90285159; PubMed=2355006;
 RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
 RA Gregory T.J.;
 RT "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in

RT Chinese hamster ovary cells.";
 RL J. Biol. Chem. 265:10373-10382 (1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M15654; AAA44205.1; -;
 DR PIR; A03973; VCLJH3
 DR HIV; M15654; ENVSEH102.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 625 625
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97224 MW; 0BF6B1A18931BB27 CRC64;

Query Match 94.9%; Score 277; DB 1; Length 856;
 Best Local Similarity 94.2%; Pred. No. 9.1e-23;
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSIEESQKQKNEQELLELDKWASLWNWFI 52
 DB 624 NNMTWMEWDREINNYTSLIHSIEESQKQKNEQELLELDKWASLWNWFI 675

RESULT 6
ENV_HV1BR STANDARD; PRT; 861 AA.
AC P03377;
21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE=85099333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV";
RL Cell 40:9-17(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02013; BAB59751.1; -;
DR EMBL; A04321; CAA00352.1; -;
DR PIR; A03975; VCLJLV.
DR PDB; 1ERF; 20-FEB-02.
DR HIV; K02013; ENV\$BRU.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 517 861 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 210 BY SIMILARITY.
FT DISULFID 126 201 BY SIMILARITY.
FT DISULFID 131 162 BY SIMILARITY.
FT DISULFID 223 252 BY SIMILARITY.
FT DISULFID 233 244 BY SIMILARITY.
FT DISULFID 301 336 BY SIMILARITY.
FT DISULFID 383 450 BY SIMILARITY.
FT DISULFID 390 423 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 642 642 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 861 AA; 97487 MW; 04DE2B4D4E4FD63A CRC64;
Query Match 94.9%; Score 277; DB 1; Length 861;
Best Local Similarity 94.2%; Pred. No. 9.2e-23;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNFNI 52
Db 629 NNTWMEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNFNI 680
RESULT 7
ENV_HV1MF STANDARD; PRT; 853 AA.
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonica C., Mann A.M., Meier C.,
RA Wasiaik A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
RT variants diminished in the ability to induce syncytium-independent
RT cytolysis";
RL J. Virol. 64:3792-3803(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33943; AAA44850.1; -;
DR PDB; 1A1K; 16-JUN-97.
DR HIV; M33943; ENV\$MFA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 31 509 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 510 853 BY SIMILARITY.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 203 BY SIMILARITY.
FT DISULFID 126 194 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 329 BY SIMILARITY.
FT DISULFID 376 443 BY SIMILARITY.

```
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 404 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 609 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 635 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 672 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 814 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 33778993B6F222ABA CRC64;

Query Match 93.8%; Score 274; DB 1; Length 853;
Best Local Similarity 92.3%; Pred. No. 1.9e-22;
Matches 48; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 52
Db 622 NNTMTWEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 673

RESULT 8
ENV_HV1PV STANDARD; PRT; 856 AA.
AC F03376;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11700;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=8511157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RT AIDS/lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02083; AAB59873.1; -.
DR EMBL; X01762; CAA25903.1; ALT_SEQ.
```

```
DR PIR; A03974; VCLJVL.
DR HIV; K02083; ENVSFV22.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT CHAIN 1 30
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 750 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 816 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C1209B3 CRC64;

Query Match 91.8%; Score 268; DB 1; Length 856;
Best Local Similarity 90.4%; Pred. No. 8.9e-22;
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 52
Db 624 NNTMTWEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 675

RESULT 9
ENV_HV1W1 STANDARD; PRT; 856 AA.
AC P31872;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).
```

OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 OX NCBI_TaxID=31678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86218077; PubMed=2423250;
 RA Starck B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
 RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
 RT "Identification and characterization of conserved and variable
 RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of
 RT AIDS.";
 RL Cell 45:637-648 (1986).
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
 CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
 CC WAS PERINATALLY INFECTED BY HER MOTHER.
 DR PIR; A24774; VCUJ3W.
 DR PDB; 1LBO; 04-DEC-02.
 DR PDB; 1LCX; 04-DEC-02.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 FT SIGNAL 1 29
 FT CHAIN 30 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 511 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 205 BY SIMILARITY.
 FT DISULFID 125 196 BY SIMILARITY.
 FT DISULFID 130 152 BY SIMILARITY.
 FT DISULFID 152 247 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 330 BY SIMILARITY.
 FT DISULFID 376 444 BY SIMILARITY.
 FT DISULFID 383 417 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97526 MW; DB68D1E49C04DE9 CRC64;
 Query Match 87.0%; Score 254; DB 1; Length 856;
 Best Local Similarity 84.6%; Pred. No. 3e-20;
 Matches 44; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NHTTWLEWDREINNTSLIHSLEESONQOEKKEQELLELDKQWASLNNWFSI 52
 Db 624 NNTTWLEWDREINNTSLIHSLEESONQOEKKEQELLELDKQWASLNNWFSI 675

RESULT 10

ENV_HVISC
 ID ENV_HVISC STANDARD; PRT; 856 AA.
 AC P05878;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88219542; PubMed=3369091;
 RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
 RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
 RT "Envelope sequences of two new United States HIV-1 isolates.";
 RL Virology 164:531-536(1988).
 CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
 CC 1984 IN SOUTHERN CALIFORNIA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M17450; -; NOT ANNOTATED_CDS.
 DR HIV; M17450; ENVSSC.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 SIGNAL.
 FT SIGNAL 1 29
 FT CHAIN 30 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 511 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT SITE 760 760 IN-FRAME TERMINATION CODON.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 206 BY SIMILARITY.
 FT DISULFID 125 197 BY SIMILARITY.
 FT DISULFID 130 160 BY SIMILARITY.
 FT DISULFID 219 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 330 BY SIMILARITY.
 FT DISULFID 376 439 BY SIMILARITY.
 FT DISULFID 383 412 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).

| | | | | | |
|---|----------|---|-----|------------------------|--------------|
| FT | CARBOHYD | 157 | 157 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 185 | 185 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 188 | 188 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 198 | 198 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 235 | 235 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 242 | 242 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 263 | 263 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 277 | 277 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 290 | 290 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 296 | 296 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 331 | 331 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 338 | 338 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 353 | 353 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 384 | 384 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 390 | 390 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 402 | 402 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 441 | 441 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 445 | 445 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 458 | 458 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 459 | 459 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 462 | 462 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 608 | 608 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 613 | 613 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 622 | 622 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 634 | 634 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 671 | 671 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| SQ | SEQUENCE | 853 AA; 97043 MW; 849B08CBAPF7008 CRC64; | | | |
| Query Match 82.5%; Score 241; DB 1; Length 853; | | | | | |
| Best Local Similarity 85.7%; Pred. No. 8.1e-19; | | | | | |
| Matches 42; Conservative 4; Mismatches 3; Indels 0; Gaps 0; | | | | | |
| Qy | 4 | TWLEWDREINNYTSLIHSLIEESQOQEKNEQELLELDKWASLWNWNI | 52 | | |
| Db | 624 | TWMEWEREIDNYTGLIYRLIEESQTQOQEKNEQELLELDKWASLWNWNI | 672 | | |
| Search completed: February 25, 2004, 15:41:44 | | | | | |
| Job time : 11.5316 secs | | | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:35:54 ; Search time 51.0127 Seconds
(without alignments)
321.625 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292

Sequence: 1 NHTWLEWREINNTSLIH.....NQELLELDKWSLWNWNI 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 290 | 99.3 | 748 | 15 Q70606 | Q70606 human immun |
| 2 | 290 | 99.3 | 752 | 15 Q70604 | Q70604 human immun |
| 3 | 290 | 99.3 | 752 | 15 Q70605 | Q70605 human immun |
| 4 | 286 | 97.9 | 747 | 15 Q70607 | Q70607 human immun |
| 5 | 286 | 97.9 | 752 | 15 Q70608 | Q70608 human immun |
| 6 | 285 | 97.6 | 856 | 15 Q90SM7 | Q90SM7 human immun |
| 7 | 282 | 96.6 | 852 | 15 Q89797 | Q89797 human immun |
| 8 | 279 | 95.5 | 856 | 15 Q92877 | Q92877 simian-huma |
| 9 | 277 | 94.9 | 645 | 15 Q993A6 | Q993A6 human immun |
| 10 | 277 | 94.9 | 851 | 15 Q78243 | Q78243 human immun |
| 11 | 277 | 94.9 | 854 | 15 Q85582 | Q85582 human immun |
| 12 | 277 | 94.9 | 854 | 15 Q72502 | Q72502 human immun |
| 13 | 277 | 94.9 | 856 | 15 Q74090 | Q74090 human immun |
| 14 | 277 | 94.9 | 856 | 15 Q74599 | Q74599 human immun |
| 15 | 273 | 93.5 | 854 | 15 Q90178 | Q90178 human immun |
| 16 | 273 | 93.5 | 854 | 15 Q78705 | Q78705 human immun |

| | | | | | |
|----|-----|------|-----|-----------|--------------------|
| 17 | 266 | 91.1 | 757 | 15 Q9Q722 | Q9Q722 human immun |
| 18 | 265 | 90.8 | 855 | 15 Q9E1R7 | Q9E1R7 human immun |
| 19 | 264 | 90.4 | 848 | 15 Q69990 | Q69990 human immun |
| 20 | 263 | 90.1 | 851 | 15 Q56110 | Q56110 human immun |
| 21 | 263 | 90.1 | 852 | 15 Q7SVL5 | Q7SVL5 human immun |
| 22 | 262 | 89.7 | 727 | 15 Q9Q723 | Q9Q723 human immun |
| 23 | 262 | 89.7 | 847 | 15 Q7ZB20 | Q7ZB20 human immun |
| 24 | 261 | 89.4 | 616 | 15 Q993B0 | Q993B0 human immun |
| 25 | 261 | 89.4 | 618 | 15 Q993B2 | Q993B2 human immun |
| 26 | 260 | 89.0 | 358 | 15 Q78120 | Q78120 human immun |
| 27 | 260 | 89.0 | 848 | 15 Q69988 | Q69988 human immun |
| 28 | 259 | 88.7 | 635 | 15 Q90U82 | Q90U82 human immun |
| 29 | 259 | 88.7 | 838 | 15 Q03806 | Q03806 human immun |
| 30 | 259 | 88.7 | 850 | 15 Q7SVL4 | Q7SVL4 human immun |
| 31 | 259 | 88.7 | 852 | 15 Q69992 | Q69992 human immun |
| 32 | 259 | 88.7 | 854 | 15 Q78225 | Q78225 human immun |
| 33 | 259 | 88.7 | 855 | 15 Q03805 | Q03805 human immun |
| 34 | 259 | 88.7 | 860 | 15 Q7SVL7 | Q7SVL7 human immun |
| 35 | 258 | 88.4 | 862 | 15 Q9E1S2 | Q9E1S2 human immun |
| 36 | 257 | 88.0 | 856 | 15 Q72993 | Q72993 human immun |
| 37 | 257 | 88.0 | 859 | 15 Q8Q863 | Q8Q863 human immun |
| 38 | 256 | 87.7 | 851 | 15 Q8Q852 | Q8Q852 human immun |
| 39 | 256 | 87.7 | 863 | 15 Q9WJU4 | Q9WJU4 human immun |
| 40 | 256 | 87.7 | 864 | 15 Q9IP48 | Q9IP48 human immun |
| 41 | 255 | 87.3 | 847 | 15 Q69996 | Q69996 human immun |
| 42 | 255 | 87.3 | 854 | 15 Q56112 | Q56112 human immun |
| 43 | 255 | 87.3 | 859 | 15 Q7SUT0 | Q7SUT0 human immun |
| 44 | 255 | 87.3 | 859 | 15 Q7SUS9 | Q7SUS9 human immun |
| 45 | 255 | 87.3 | 863 | 15 Q9WJU8 | Q9WJU8 human immun |

ALIGNMENTS

RESULT 1

Q70606 PRELIMINARY; PRT; 748 AA.
ID Q70606;
AC Q70606;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]_TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=LW881;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.,
RT "Viral variability and serum antibody response in a laboratory worker
infected with HIV type 1 (HTLV type IIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LW881;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12032; AAA76668.1; -;
DR PIR; A53591; A53591.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0018028; C:iviral capsid; IEA.
DR GO; GO:0019033; C:iviral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 748 748
SQ SEQUENCE 748 AA; 84224 MW; 56BEDF186C67694B CRC64;

SQ SEQUENCE 747 AA; 84250 MW; 732E836A52245F14 CRC64;

Query Match 97.9%; Score 286; DB 15; Length 747;
Best Local Similarity 98.1%; Pred. No. 8.9e-23;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWNI 52
DB 619 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWNI 670

RESULT 5

ID Q70608 PRELIMINARY; PRT; 752 AA.
AC Q70608;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LW87-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
infected with HIV type 1 (HTLV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LW87-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12035; AAA76670.1; -;
DR PIR; A53591; A53591.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
FT NON TER 752 752
SQ SEQUENCE 752 AA; 84780 MW; 708672A2D0C0E8F8 CRC64;

Query Match 97.9%; Score 286; DB 15; Length 752;
Best Local Similarity 98.1%; Pred. No. 9e-23;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWNI 52
DB 624 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWNI 675

RESULT 6

ID Q90SM7 PRELIMINARY; PRT; 856 AA.
AC Q90SM7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=HXB2;
RA Ataman-Onal Y., Cheynet V., Verrier B.;
RT "Mutations and transcriptional alterations associated with the
downregulation of HIV-1 envelope glycoprotein expression following
acute cytopathic effects";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF358141; AAK49977.1; -;
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 856 AA; 97126 MW; 9458D02B2FD734B3 CRC64;

Query Match 97.6%; Score 285; DB 15; Length 856;
Best Local Similarity 96.2%; Pred. No. 1.3e-22;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWNI 52
DB 624 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWNI 675

RESULT 7

ID Q89797 PRELIMINARY; PRT; 852 AA.
AC Q89797;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LW90-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
infected with HIV type 1 (HTLV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LW90-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12053; AAA76685.1; -;
DR EMBL; U12036; AAA76671.1; -;
DR PIR; A53591; A53591.
DR PDB; 1F23; 20-JUN-01.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polypeptide;
Transmembrane.
SQ SEQUENCE 852 AA; 96347 MW; 4E865229D4EB33CF CRC64;

Query Match 96.6%; Score 282; DB 15; Length 852;

```

Best Local Similarity 96.2%; Pred. No. 2.8e-22; Mismatches 1; Indels 0; Gaps 0;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQOQKNEQELLELDKQWASLWNNWFI 52
   |||:|||||
Db 620 NRTTWMEWDREINNTSLIHSLEESQOQKNEQELLELDKQWASLWNNWFI 671

RESULT 8
O92877 ID O92877 PRELIMINARY; PRT; 856 AA.
AC O92877;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Simian-Human immunodeficiency virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=57667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098984; PubMed=9882298;
RA Cayabyab M., Karlsson G.B., Etemad-Moghadam B.A., Hofmann W.,
RA Halloran M., Axthelm M.W., Letvin N.L., Sodroski J.G.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL; AF041850; AAD12142.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR EnvPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 856 AA; 97151 MW; C50BE0388F73659 CRC64;

Query Match 95.5%; Score 279; DB 15; Length 856;
Best Local Similarity 94.2%; Pred. No. 5.9e-22;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQOQKNEQELLELDKQWASLWNNWFI 52
   |||:|||||
Db 624 NHTTWMEWDREINNTSLIHSLEESQOQKNEQELLELDKQWASLWNNWFI 675

RESULT 9
Q993A6 ID Q993A6 PRELIMINARY; PRT; 645 AA.
AC Q993A6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Truncated envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1007;
RC MEDLINE=21192672; PubMed=11287644;
RA Surman S., Lockey T.D., Slobod K.S., Jones B., Riberty J.M.,

```

```

RA White S.W., Doherty P.C., Hurwitz J.L.;
RT "Localization of CD4+ T cell epitope hotspots to exposed strands of
RT HIV envelope glycoprotein suggests structural influences on antigen
RT processing.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4587-4592(2001).
DR EMBL; AF315153; AAK18810.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polypeptide;
KW Transmembrane.
FT NON_TER 1
SQ SEQUENCE 645 AA; 72485 MW; B076514BE93362EC CRC64;

Query Match 94.9%; Score 277; DB 15; Length 645;
Best Local Similarity 94.2%; Pred. No. 7.2e-22;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQOQKNEQELLELDKQWASLWNNWFI 52
   |||:|||||
Db 593 NNTTWMEWDREINNTSLIHSLEESQOQKNEQELLELDKQWASLWNNWFI 644

RESULT 10
Q78243 ID Q78243 PRELIMINARY; PRT; 851 AA.
AC Q78243;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Carlini F., Federic M., Equestre M., Ricci S., Ratti G., Zibai Q.,
RA Verani P., Rossi G.B.;
RT "Sequence analysis of HIV-1 proviral DNA from a non producer
RT chronically infected HUT-78 cellular clone.";
RL J. Viral Diseases 1:40-55(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352106; PubMed=2765297;
RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,
RA Macchi B., Mangiano N., Verani P., Rossi G.;
RT "Biological and molecular characterization of producer and non
RT producer clones from HUT-78 infected with a patient HIV isolate.";
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,
RA Borsetti A., Saggio I., Verani P., Rossi G.;
RT "Variability of HIV-1 virus: characteristics of an infected but not
RT productive clone.";
RL Int. J. Immunopharmacol. 3:17-23(1990).
DR EMBL; Z11530; CAA77628.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR PIR; S33985; S33985.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:09 ; Search time 125.348 Seconds
(without alignments)
191.599 Million cell updates/sec

Title: US-09-877-606-4
Perfect score: 452
Sequence: 1 NHTWLEWDREINNYTSLIH.....SKYHIENEIARIKKLIGER 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 373.5 | 82.6 | 146 | 6 | ABP98163 Amino aci |
| 2 | 314.5 | 69.6 | 185 | 6 | ABb84628 gp41 ecto |
| 3 | 290 | 64.2 | 139 | 6 | ABb84625 HIV-1 str |
| 4 | 290 | 64.2 | 139 | 6 | ABb84626 HIV-1 str |
| 5 | 290 | 64.2 | 268 | 2 | AAY22820 SEQ ID NO |
| 6 | 290 | 64.2 | 268 | 5 | ABG68291 Envelope |
| 7 | 290 | 64.2 | 268 | 6 | ABU57698 Human imm |
| 8 | 290 | 64.2 | 344 | 5 | ABb83400 HIV gp41 |
| 9 | 290 | 64.2 | 345 | 3 | ADc14130 HIV gp41 |
| 10 | 290 | 64.2 | 346 | 7 | ABb80284 HIV gp41 |
| 11 | 290 | 64.2 | 359 | 6 | ABr57594 HIV gp41 |
| 12 | 290 | 64.2 | 420 | 2 | AAR53785 Translati |
| 13 | 290 | 64.2 | 519 | 2 | AAW00181 HIV-1 env |
| 14 | 290 | 64.2 | 521 | 2 | AAR69997 HIV-1 env |
| 15 | 290 | 64.2 | 853 | 2 | AAW43066 HIV-1 gpl |
| 16 | 290 | 64.2 | 856 | 1 | AAp60131 Sequence |
| 17 | 290 | 64.2 | 856 | 2 | AAR41025 Selective |
| 18 | 290 | 64.2 | 856 | 2 | AAR41032 Selective |
| 19 | 290 | 64.2 | 856 | 2 | AAR41031 Selective |
| 20 | 290 | 64.2 | 856 | 2 | AAR41026 Selective |
| 21 | 290 | 64.2 | 856 | 2 | AAR41028 Selective |
| 22 | 290 | 64.2 | 856 | 2 | AAR41029 Selective |
| 23 | 290 | 64.2 | 856 | 2 | AAR41030 Selective |
| 24 | 290 | 64.2 | 856 | 2 | AAR41027 Selective |
| 25 | 290 | 64.2 | 856 | 3 | AAY97072 Wild type |

| | | | | | |
|----|-----|------|-----|---|--------------------|
| 26 | 290 | 64.2 | 856 | 4 | AAB45697 HIV-1/III |
| 27 | 290 | 64.2 | 856 | 6 | ABU63327 Human lym |
| 28 | 290 | 64.2 | 856 | 7 | ABb80283 gp41 of 1 |
| 29 | 290 | 64.2 | 856 | 8 | ADe84721 Human imm |
| 30 | 278 | 61.5 | 177 | 5 | ABb83401 HIV gp41 |
| 31 | 278 | 61.5 | 177 | 5 | ABb83413 HIV gp41 |
| 32 | 278 | 61.5 | 200 | 7 | ADc24667 HIV gp41 |
| 33 | 277 | 61.3 | 56 | 3 | ABb14532 HIV-1 iso |
| 34 | 277 | 61.3 | 56 | 4 | AAG63860 Amino aci |
| 35 | 277 | 61.3 | 56 | 4 | AAU70191 HIV viral |
| 36 | 277 | 61.3 | 138 | 3 | AAV69792 HIV-1 gp4 |
| 37 | 277 | 61.3 | 147 | 6 | ABP71557 HIV-1 env |
| 38 | 277 | 61.3 | 150 | 1 | AAp80745 Sequence |
| 39 | 277 | 61.3 | 192 | 2 | AAy24098 Recombina |
| 40 | 277 | 61.3 | 233 | 2 | AAr93199 Recombina |
| 41 | 277 | 61.3 | 268 | 2 | AAy22822 SEQ ID NO |
| 42 | 277 | 61.3 | 268 | 2 | AAy22821 SEQ ID NO |
| 43 | 277 | 61.3 | 268 | 5 | ABG68292 Envelope |
| 44 | 277 | 61.3 | 268 | 5 | ABG68293 Envelope |
| 45 | 277 | 61.3 | 268 | 6 | ABU57699 Human imm |

ALIGNMENTS

RESULT 1
ABP98163
ID ABP98163 standard; peptide; 146 AA.
XX AC ABP98163;
XX DT 11-AUG-2003 (first entry)
XX DE Amino acid sequence of the GCG2 fusion protein.
XX KW Immunogen; pathogenesis factor.
XX OS Synthetic.
XX PN WC2003017032-A2.
XX PD 27-FEB-2003.
XX PF 08-AUG-2002; 2002WO-US025263.
XX PR 14-AUG-2001; 2001US-0312276P.
XX PA (DAND) DANA FARBER CANCER INST INC.
XX PI Reinherz EL, Kim M, Reche PA, Wang J;
XX DR WPI; 2003-278593/27.
XX PT New computer-assisted method of generating an immunogen by providing a
XX PT pathogenesis factor (PI;) or its fragment and determining whether the
XX PS produced compound is an immunogen in a mammalian host.
XX PS Disclosure; Fig 4A; 53pp; English.
XX CC The specification describes a method for the computer-assisted generation
XX CC of an immunogen. The method requires use of a programmed computer
XX CC comprising a processor and an input device comprising providing a
XX CC pathogenesis factor or its fragment comprising a region with low
XX CC polymorphism, and determining whether the produced compound is an
XX CC immunogen in a mammalian host. The method is useful for generating an
XX CC immunogen. The present sequence represents the GCG2 fusion protein. It is
XX CC used in the course of the invention
SQ Sequence 146 AA;

Query Match 82.6%; Score 373.5; DB 6; Length 146;
Best Local Similarity 79.8%; Pred. No. 9.4e-29;
Matches 71; Conservative 8; Mismatches 3; Indels 7; Gaps 1;

CC administered to the body of a patient as a primary immunisation, then
 CC mucosally administered as a secondary immunisation. This sequence
 CC represents the HIV-1 strain HXB2 native envelope glycoprotein gp41
 CC corresponding to residues 29-167 of the ectodomain. This native protein
 CC is modified to eliminate a disulphide bond and many residues are
 CC deglycosylated prior to the construction of the fusion protein
 CC represented in ABB84628. (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 139 AA;
 Query Match 64.2%; Score 290; DB 6; Length 139;
 Best Local Similarity 98.1%; Pred. No. 1.2e-20;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWFI 52
 Db 85 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWFI 136
 RESULT 4
 ABB84626
 ID ABB84626 standard; protein; 139 AA.
 AC ABB84626;
 XX
 DT 04-FEB-2003 (first entry)
 DE HIV-1 strain HXB2 mutant gp41 ectodomain fragment.
 KW Envelope glycoprotein; gp41; fusion protein; haemagglutinin protein;
 KW immunogen; vaccine; IGG; IGA; immunisation; mutant; muten.
 XX
 OS Human immunodeficiency virus 1.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 59 /label= C59S
 FT /note= "Wild type Cys is replaced by Ser"
 FT Misc-difference 65 /label= C65S
 FT /note= "Wild type Cys is replaced by Ser"
 FT Misc-difference 125 /note= "This residue is represented as Asp in the
 FT specification but is described as Asp the corresponding
 FT fusion protein represented in ABB84628"
 XX
 PN WO200281655-A2.
 XX
 PD 17-OCT-2002.
 XX
 PP 27-MAR-2002; 2002WO-US009353.
 XX
 PR 28-MAR-2001; 2001US-0279383P.
 XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 XX
 XX Weissenhorn W, Wiley D, Mantis N, Neutra MR, Kozlowski P;
 XX WPI; 2003-058532/05.
 XX
 PT Novel fusion protein for inducing human immunodeficiency virus-antigen
 PT specific IGG and IGA antibodies, has ectodomain of HIV-1 envelope
 PT glycoprotein gp41 fused to fragment of influenza virus haemagglutinin
 PT protein.
 XX
 PS Disclosure; Page; 78pp; English.
 XX
 CC This invention describes a novel fusion protein construct comprising a
 CC fragment of the human immunodeficiency virus envelope glycoprotein gp41
 CC ectodomain and a second amino acid region composed of a fragment of the
 CC influenza virus haemagglutinin protein. The composition of the invention
 CC is soluble at physiological pH and is useful as an immunogen, useful in a

CC vaccine for the induction of human immunodeficiency virus (HIV)-antigen
 CC specific serum IGG and secretory IGA antibodies in vivo. The immunogen is
 CC administered to the body of a patient as a primary immunisation, then
 CC mucosally administered as a secondary immunisation. This sequence
 CC represents a HIV-1 strain HXB2 mutant envelope glycoprotein gp41
 CC corresponding to residues 29-167 of the ectodomain, which is used in the
 CC construction of the fusion protein represented in ABB84628. The wild-type
 CC protein has been modified to eliminate a disulphide bond and many
 CC residues have been deglycosylated to generate the mutant. NOTE: This
 CC sequence does not appear in the disclosure and has been created using the
 CC wild-type protein represented in ABB84625 and from information given on
 CC Page 26 of the Disclosure
 XX
 SQ Sequence 139 AA;
 Query Match 64.2%; Score 290; DB 6; Length 139;
 Best Local Similarity 98.1%; Pred. No. 1.2e-20;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWFI 52
 Db 85 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWASLWNWFI 136
 RESULT 5
 AAY22820
 ID AAY22820 standard; protein; 268 AA.
 AC AAY22820;
 XX
 DT 19-AUG-1999 (first entry)
 DE SEQ ID NO. 16 from WO9820036.
 XX
 KW HIV; gp41 protein; constrained helical peptide; HIV infection; vaccine;
 KW antibody; viral membrane fusion; viral infectivity;
 KW ligand affinity purification; protein A replacement;
 KW immunoglobulin purification; epitope mimic.
 XX
 OS Human immunodeficiency virus.
 XX
 PN WO9820036-A1.
 XX
 PD 14-MAY-1998.
 XX
 PP 05-NOV-1997; 97WO-US020069.
 XX
 PR 06-NOV-1996; 96US-00743698.
 PR 16-JUN-1997; 97US-00876698.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Braisted A, Judice JK, McDowell RS, Phelan JC, Starovaasnik MA;
 PI Wells JA;
 XX
 XX WPI; 1998-286866/25.
 XX
 PT Production of constrained helical peptide(s) by linking side chains on
 PT termini of octa-peptide - derived from human immunodeficiency virus gp41
 PT protein, useful in vaccines for treatment and prevention of infection.
 XX
 PS Claim 11; Page 156; 279pp; English.
 XX
 CC Peptides AAY22805-Y22917 are derived from Human immunodeficiency virus
 CC (HIV). Specifically, AAY22810-Y22910 are derived from gp41 proteins of
 CC known HIV virus strains (AAY22810, AAY22871, AAY22880, AAY22888 and
 CC AAY22903 represent consensus sequences of various sections of the gp41
 CC protein). Sequences derived from the peptides are used to produce
 CC constrained helical peptides of the invention. The constrained helical
 CC peptide is produced by synthesising an octapeptide in which both terminal
 CC amino acids have a side-chain that includes a group able to form an amide
 CC bond, and cyclising the octapeptide by reacting the specified side-chain
 CC residues with a difunctional linker to produce two amide bonds. The

CC constrained helical peptides are used to treat or prevent HIV infection,
CC especially as vaccines that generate antibodies that prevent viral
CC membrane fusion or infectivity. Vaccines may contain constrained helical
CC peptides derived from several different strains of HIV. The antibodies
CC are also useful for diagnosing HIV infection. Other uses for the
CC constrained helical peptides are in affinity purification of ligands
CC (particularly where complete binding protein is not readily available,
CC e.g. replacements for protein A in immunoglobulin purification); as
CC epitope mimics for antibody production; for isolation of synthetic
CC antibody clones from phage display libraries, or as stable forms of
CC "floppy" peptides or proteins

XX
SQ Sequence 268 AA;

Query Match 64.2%; Score 290; DB 2; Length 268;

Best Local Similarity 98.1%; Pred. No. 2.6e-20;

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNWFI 52

Db 154 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNWFI 205

RESULT 6

ABG68291

ID ABG68291 standard; protein; 268 AA.

XX

AC ABG68291;

XX

DT 29-AUG-2003 (revised)

DT 07-OCT-2002 (first entry)

XX

DE Envelope protein gp41 from HIV clade B strain #10.

XX

KW HIV; glycoprotein; gp41; antigen; helical conformation;

KW virus-induced membrane fusion; acquired immunodeficiency syndrome; AIDS;

KW viral envelope protein; vaccine; virucide; anti-HIV.

XX

OS Human immunodeficiency virus 1; clade B.

XX

PN US6271198-B1.

XX

PD 07-AUG-2001.

XX

PF 05-NOV-1997; 97US-00965056.

XX

PR 06-NOV-1996; 96US-00743698.

XX

PR 16-JUN-1997; 97US-0049787P.

XX

PR 16-JUN-1997; 97US-00876698.

XX

PA (GETH) GENENTECH INC.

XX

PI Braisted AC, Judice JK, McDowell RS, Phelan JC, Starovasnik MA;

XX

PI Wells JA;

XX

DR WPI; 2002-487624/52.

XX

XX

PT New cyclic peptides from human immune deficiency virus gp41, useful for

PT treatment or prevention of HIV infection, are constrained to have alpha-

PT helical conformation.

XX

PS Disclosure; Col 151-154; 175pp; English.

XX

CC The invention relates to cyclic peptides (A) with a constrained helical
CC conformation, derived from gp41 (glycoprotein 41, a viral envelope
CC protein) of human immunodeficiency virus (HIV). The cyclic
CC peptides have formulas given in the specification part of which are
CC derived from a consensus sequence of gp41 derived from HIV clades A, B,
CC C, D, E or O. The peptides are used to cause induction of a specific
CC immune response, resulting in antibodies that prevent virus-induced
CC membrane fusion. The peptides are used to treat subjects with, or at risk
CC of, HIV infection, either as antifusion/anti-infection agents or,
CC preferably where associated with a carrier, as an immunogen (including as

CC vaccine) to raise antibodies. The antibodies may be used for diagnosis or
CC prevention/treatment of HIV infection (i.e. acquired immunodeficiency
CC syndrome, AIDS), e.g. prevention of mother-to-child transmission or in
CC cases of health care accidents. The peptides can be based on specific HIV
CC strains, e.g. breakthrough isolates of HIV that have developed during
CC vaccine trials, so a combination of them should cover a wide range of
CC protection. The present sequence is gp41 protein from a particular HIV
CC clade used to derive a consensus sequence of gp41. (Updated on 29-AUG-
CC 2003 to standardise OS field)

XX
SQ Sequence 268 AA;

Query Match 64.2%; Score 290; DB 5; Length 268;

Best Local Similarity 98.1%; Pred. No. 2.6e-20;

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNWFI 52

Db 154 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKWSLWNWFI 205

RESULT 7

ABU57698

ID ABU57698 standard; protein; 268 AA.

XX

AC ABU57698;

XX

DT 10-APR-2003 (first entry)

XX

DE Human immunodeficiency virus (HIV) envelope protein gp41 #10.

XX

KW Human immunodeficiency virus; HIV; vaccine; helical peptide compound;

KW viral membrane fusion; haptens; immunogen; peptidomimetic; gp41;

XX

OS Human immunodeficiency virus.

XX

PN US2002151473-A1.

XX

PD 17-OCT-2002.

XX

PF 15-MAY-2001; 2001US-00854816.

XX

PR 06-NOV-1996; 96US-00743698.

XX

PR 16-JUN-1997; 97US-0049787P.

XX

PR 16-JUN-1997; 97US-00876698.

XX

XX

PA (BRAI/) BRAISTED A C.

XX

PA (JUDI/) JUDICE J K.

XX

PA (MCDO/) MCDOWELL R S.

XX

PA (PHEL/) PHELAN J C.

XX

PA (STAR/) STAROVASNIK M A.

XX

PA (WELL/) WELLS J A.

XX

PI Braisted AC, Judice JK, McDowell RS, Phelan JC, Starovasnik MA;

XX

PI Wells JA;

XX

DR WPI; 2003-182525/18.

XX

XX

PT Novel constrained helical peptide compound useful for prophylactically or

PT therapeutically treating mammal at risk for or infected with human

PT immunodeficiency virus.

XX

PS Disclosure; Fig 16; 180pp; English.

XX

CC The invention describes a constrained helical peptide compound (I)
CC comprising a first constrained helical peptide comprising a sequence of 8
CC amino acids (a.a.s) having a first and second terminal residue both
CC flanking an internal sequence of 6 a.a.s, where the terminal residues have
CC a side chain that are linked to each other forming a locking group to
CC form a constrained helical peptide. (I) is useful for preparing
CC antibodies that prevent viral membrane fusion, as haptens, preferably

CC attached to a carrier, for use as an immunogen to raise antibodies that
 CC have a diagnostic use, as a vaccine for treatment of patients at risk of
 CC or infected with HIV, to create combinatorial constrained helical peptide
 CC libraries that are useful in chemical selection systems, to isolate the
 CC binding determinants from alpha-helical binding domains of known
 CC proteins, for determining whether a binding determinate in an alpha-
 CC helical binding domain of a known protein can serve as a structural model
 CC for the design of peptidomimetics, to replace intact binding proteins or
 CC protein binding domains in the affinity purification of ligands, to mimic
 CC epitopes in proteins to selectively raise polyclonal or monoclonal
 CC antibodies against such individual epitopes for isolating synthetic
 CC antibody clones with a selected binding activity from phage display
 CC combinatorial libraries, to provide conformationally stable variants of
 CC peptides or proteins which exhibit floppy or unstable alpha-helical
 CC secondary structure at one or more sites in unrestrained form under
 CC conditions of interest. This is the amino acid sequence of an HIV
 CC envelope protein gp41, fragments of which are used in the creation of
 CC locked helix peptides

XX SQ Sequence 268 AA;

Query Match 64.2%; Score 290; DB 6; Length 268;
 Best Local Similarity 98.1%; Pred. No. 2.6e-20;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQELLELDKWASLWNWFNI 52
 |||||:|||||
 DB 154 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQELLELDKWASLWNWFNI 205

RESULT 8

ABB83400
 ID . ABB83400 standard; protein; 344 AA.

XX AC ABB83400;

XX DT 19-SEP-2002 (first entry)

XX DE HIV gp41 LAI protein #1.

XX KW HIV; gp41; LAI; virucide; anti-HIV; vaccine; HIV infection.

XX OS Human immunodeficiency virus.

XX FH Key Location/Qualifiers

XX FT Peptide 1..23

XX FT Domain /label= Fusion_peptide

XX FT 173..194

XX FT /label= Transmembrane_domain

XX PN WO200253587-A2.

XX PD 11-JUL-2002.

XX PF 04-JAN-2002; 2002WO-FR000031.

XX PR 05-JAN-2001; 2001PR-00000141.

XX PR 23-JAN-2001; 2001PR-00000848.

XX PA (AVET) AVENTIS PASTEUR.

XX PI Brasseur R, Charlotiaux B, Chevalier M, El Habib R, Krell T;

XX PI Sodoyer R;

XX XX WPI; 2002-528852/56.

XX XX New mutant human immune deficiency virus gp41 polypeptide, useful in
 PT vaccines against human immune deficiency virus, mimics the intermediate
 PT state of the native polypeptide.

XX PS Disclosure; Page 29; 29pp; French.

XX CC The present invention relates to novel mutant HIV gp41 LAI proteins

CC (ABB83411-ABB83420). The muteins can form a structure corresponding to,
 CC or mimicking, the intermediate state of gp41. The muteins, or their
 CC conjugate with a carrier, or vectors containing nucleic acid that encode
 CC them, are used in vaccines for treatment or prevention of infection by
 CC HIV. The present sequence is wild-type HIV gp41 LAI protein, which was
 CC used to generate the muteins of the invention

XX SQ Sequence 344 AA;

Query Match 64.2%; Score 290; DB 5; Length 344;
 Best Local Similarity 98.1%; Pred. No. 3.5e-20;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQELLELDKWASLWNWFNI 52
 |||||:|||||

DB 113 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQELLELDKWASLWNWFNI 164

RESULT 9

ADC14130

XX ID ADC14130 standard; protein; 345 AA.

XX AC ADC14130;

XX DT 18-DEC-2003 (first entry)

XX DE HIV gp41 protein used to create the variants Trx-N and GST-C.

XX KW HIV infection; gp41; immunoassay; Trx-N; thioredoxin; GST-C;

XX KW glutathione-S-transferase; AIDS; HIV.

XX OS Human immunodeficiency virus 1.

XX PN US6294341-B1.

XX PD 25-SEP-2001.

XX PF 19-MAR-1999; 99US-00272342.

XX PR 21-MAR-1998; 98KR-00009858.

XX PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.

XX PI Yu YG, Kim S, Ryu J;

XX DR WPI; 2000-585178/73.

XX PT Detecting inhibitors of gp41 activity, useful for detecting inhibitors of
 PT human immunodeficiency virus (HIV) infection, by employing an assay
 PT detecting variant protein thioredoxin-N and the glutathione S-transferase
 PT protein interaction.

XX PS Disclosure; SEQ ID NO 7; 13pp; English.

XX CC This invention relates to a novel method for detecting the presence of a
 CC substance that inhibits HIV infection. Specifically, it refers to
 CC inhibiting activity of the HIV transmembrane protein gp41, which is
 CC responsible for the infection of HIV by causing fusion between the viral
 CC envelope membrane the cell cytoplasmic membrane. The present invention
 CC describes an immunoassay that can identify the presence of an interaction
 CC between the two helical domains of gp41 by determining the interaction of
 CC two variant proteins. These variant proteins being Trx-N (N-terminal
 CC helical domain of gp41 connected to thioredoxin) and GST-C (C-terminal
 CC helical domain of gp41 linked to glutathione-S-transferase). As such, the
 CC immunoassay can be used to detect substances that bind to one of the two
 CC helical domains of gp41 thereby inhibiting gp41 activity and HIV
 CC infection, and in turn can be used to develop a therapeutic agent for
 CC AIDS. This polypeptide sequence is the HIV gp41 protein, used to create
 CC the chimeric trx-N and GST-C variant proteins of the invention.

XX SQ Sequence 345 AA;

Query Match 64.2%; Score 290; DB 3; Length 345;

```

Best Local Similarity 98.1%; Pred. No. 3.5e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 51; Conservative 1;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNWFI 52
DB 113 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNWFI 164

RESULT 10
ABB80284
ID ABB80284 standard; protein; 346 AA.
XX
AC ABB80284;
XX
DT 15-JAN-2004 (first entry)
XX
DE 511-856 of gp41 of 1987 HIV1 isolate.
XX
KW Envelope glycoprotein; gp41; 1987 HIV1; antigen; immunocomplex;
KW detection.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO2003073992-A2.
XX
PD 12-SEP-2003.
XX
PF 26-FEB-2003; 2003WO-US006206.
XX
PR 28-FEB-2002; 2002US-0360448P.
PR 18-APR-2002; 2002US-0373448P.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Cloyd MW, Chen J;
XX
XX WPI; 2003-756752/71.
DR
XX
PT Screening for HIV in a subject with a composition comprising a
PT recombinant non-denatured HIV gp41 or gp160 antigen and detecting an
PT immunocomplex between an antibody and the antigen.
XX
PS Claim 13; Page 96-97; 99pp; English.
XX
CC This sequence represents a fragment of the envelope glycoprotein, gp41,
CC of 1987 HIV1 isolate. Antigenic fragments of the gp41 protein may be used
CC in the method of the invention for screening for human immuno- deficiency
CC virus (HIV) in a subject. The method comprises contacting a sample from
CC the subject with a composition comprising a recombinant, non-denatured
CC HIV gp41 antigen under conditions that permit formation of an
CC immunocomplex between any antibody in the sample that can specifically
CC bind to the gp41 antigen; and detecting whether an immunocomplex is
CC formed between an antibody and the gp41 antigen
XX
SQ Sequence 346 AA;

Query Match 64.2%; Score 290; DB 7; Length 346;
Best Local Similarity 98.1%; Pred. No. 3.5e-20;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNWFI 52
DB 114 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNWFI 165

RESULT 11
ABR57594
ID ABR57594 standard; protein; 359 AA.
XX
AC ABR57594;
XX
DT 12-AUG-2003 (first entry)
XX
XX

```

```

DE HIV gp41 HXB2D protein SEQ ID 58.
XX
KW HIV; treatment; therapy; gp41.
XX
OS Human immunodeficiency virus.
XX
PN EP1283272-A2.
XX
PD 12-FEB-2003.
XX
PF 08-AUG-2002; 2002EP-00078298.
XX
PR 08-AUG-2001; 2001EP-00203011.
PR 08-AUG-2001; 2001US-0310497P.
XX
PA (TIBO-) TIBOTEC PHARM LTD.
XX
PI Kemp S, Vingerhoets JHJ, Michiels LEJ;
XX
XX WPI; 2003-364991/35.
DR
XX
PT Determining the susceptibility of the HIV virus to a drug by monitoring
PT molecular events at the HIV envelope protein, useful for the diagnosis,
PT evaluation of treatment and drug screening and/or drug development in HIV
PT disease.
XX
PS Example 6; Fig 8; 54pp; English.
XX
CC The present invention relates to a method for determining the
CC susceptibility of HIV to a drug. The method comprises obtaining a sample
CC comprising HIV RNA or DNA, reverse-transcribing and amplifying the RNA or
CC DNA, homologously recombining or ligating at least one amplicon with to
CC prepare a recombinant virus, and monitoring the recombinant virus in the
CC presence of the drug to determine the phenotypic susceptibility. The
CC methods and compositions of the present invention are useful for the
CC evaluation of HIV treatment, in particular for the determination of
CC molecular events at the HIV envelope protein and their effect on
CC therapeutic efficacy of drugs. The methods may find use in multiple
CC fields including diagnostics, drug screening, pharmacogenetics and drug
CC development in HIV disease. The present gp41 protein was used to
CC illustrate the method of the invention
XX
SQ Sequence 359 AA;

Query Match 64.2%; Score 290; DB 6; Length 359;
Best Local Similarity 98.1%; Pred. No. 3.6e-20;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNWFI 52
DB 135 NHTTWLEWDREINNYTSLIHSLSIESQKQKNEQLLELDKWSLWNNWFI 186

RESULT 12
AAR53785
ID AAR53785 standard; protein; 420 AA.
XX
AC AAR53785;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-DEC-1994 (first entry)
XX
DE Translation of HIV-1 in the region encoding the gp41 polypeptide.
XX
KW Human immunodeficiency virus; HIV-1; AIDS; lentivirus; TM protein;
KW extracellular protein; transmembrane protein; gp41.
XX
OS Human immunodeficiency virus 1.
XX
FH Key 43.387
FT Protein /label= gp41
FT

```

```

XX WO9412533-A1.
XX
XX PD 09-JUN-1994.
XX
XX PF 12-JAN-1993; 93WO-US0000212.
XX
XX PR 23-NOV-1992; 92US-00979975.
XX
XX PA (HARD ) HARVARD COLLEGE.
XX
XX PI Essex ME, Yu X, Lee TH;
XX
XX DR WPI; 1994-200197/24.
XX
XX DR N-PSDB; AAQ66275.
XX
XX PT Method for treating HIV patients - comprises administration of mutated
XX GP41 polypeptide.
XX
XX PS Claim 11; Fig 1; 54pp; English.
XX
XX CC The inventors claim a method of treating a patient infected with HIV by
XX administering a mutated gp41 polypeptide or a therapeutic composition
XX comprising nucleic acid encoding the mutant gp41 polypeptide in an
XX expressible genetic construction. The mutant gp41 polypeptide contains a
XX deletion of at least one AA in at least one of the following regions of
XX wild type gp41 (AAR53783): AAs 844-856; 814-856; 796-856; 776-856; 753-
XX 856; or 710-856, effective to either disrupt viral replication or HIV or
XX disrupt the assembly of viral Env proteins in an HIV infected cell.
XX AAQ66275 corresp. to bps 7631- 8890 of wt HIV-1. X in the AA sequence
XX represents the posn. of a stop codon in AAQ66275. (Updated on 25-MAR-2003
XX to correct FN field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 420 AA;

Query Match 64.2%; Score 290; DB 2; Length 420;
Best Local Similarity 98.1%; Pred. No. 4.3e-20;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQNOQEKNEQELLELDKWASLWNWFNI 52
   |||||:|||||
Db 155 NHTTWLEWDREINNTSLIHSLEESQNOQEKNEQELLELDKWASLWNWFNI 206

RESULT 13
AAW00181
ID AAW00181 standard; protein; 519 AA.
XX
XX AC AAW00181;
XX
XX DT 16-OCT-2003 (revised)
XX
XX DT 12-NOV-1996 (first entry)
XX
XX DE HIV-1 env protein.
XX
XX KW Poliovirus; encapsidation; capsid; vaccine; genetic immunisation; HIV-1;
XX human immunodeficiency virus type 1; env protein.
XX
XX OS Human immunodeficiency virus 1.
XX
XX PN WO9625173-A1.
XX
XX PD 22-AUG-1996.
XX
XX PF 13-FEB-1996; 96WO-US001895.
XX
XX PR 15-FEB-1995; 95US-00389459.
XX
XX PA (UABR-) UAB RES FOUND.
XX
XX PI Morrow CD, Porter DC, Ansardi DC;
XX
XX DR WPI; 1996-393136/39.

```

```

DR N-PSDB; AAT33295.
XX
XX CC Encapsidation of recombinant polio:virus nucleic acid for use in vaccines
XX PT - using a polio:virus nucleic acid which lacks the P1 capsid region and
XX PT an expression system which provides the region.
XX
XX PS Disclosure; Page 61-63; 108pp; English.
XX
XX CC cDNA sequences (AAT33293-95) respectively code for the gag, pol and env
XX proteins (AAW00179-81) of HIV-1. They can be used to substitute the P1
XX capsid gene of poliovirus in recombinant poliovirus nucleic acids
XX (rPNAs). Such rPNAs are encapsidated by introduction into a host cell
XX together with a vaccinia virus or plasmid vector encoding the poliovirus
XX P1 capsid precursor protein. Encapsidated rPNAs are useful for genetic
XX immunisation, stimulating an immune response to the HIV-1 protein.
XX (Updated on 16-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 519 AA;

Query Match 64.2%; Score 290; DB 2; Length 519;
Best Local Similarity 98.1%; Pred. No. 5.5e-20;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQNOQEKNEQELLELDKWASLWNWFNI 52
   |||||:|||||
Db 420 NHTTWLEWDREINNTSLIHSLEESQNOQEKNEQELLELDKWASLWNWFNI 471

RESULT 14
AAR69997
ID AAR69997 standard; protein; 521 AA.
XX
XX AC AAR69997;
XX
XX DT 16-OCT-2003 (revised)
XX
XX DT 13-SEP-1995 (first entry)
XX
XX DE HIV-1 env protein.
XX
XX KW Encapsulation; Human immunodeficiency virus; type 1; immunogenic;
XX poliovirus.
XX
XX OS Human immunodeficiency virus 1.
XX
XX PN CA2125344-A.
XX
XX PD 02-JAN-1995.
XX
XX PF 07-JUN-1994; 94CA-02125344.
XX
XX PR 01-JUL-1993; 93US-00087009.
XX
XX PA (UABR-) UAB RES FOUND.
XX
XX PI Morrow CD;
XX
XX DR WPI; 1995-099021/14.
XX
XX DR N-PSDB; AAQ80575.
XX
XX PT Method for encapsidating recombinant polio:virus nucleic acid - useful
XX for providing compns. to stimulate immune response.
XX
XX PS Disclosure; Page 38; 62pp; English.
XX
XX CC The sequence is that of the HIV-1 env protein. The DNA encoding such
XX protein is used in a method (claimed) to encapsulate poliovirus cDNA to
XX make it more immunogenic. See also AAR69995-6. (Updated on 16-OCT-2003 to
XX standardise OS field)
XX
XX SQ Sequence 521 AA;

Query Match 64.2%; Score 290; DB 2; Length 521;
Best Local Similarity 98.1%; Pred. No. 5.6e-20;

```

Search completed: February 25, 2004, 15:41:04
Job time : 126.348 secs

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NHTTWLEWREINNYTSLIHSLEESQKQKNEQLLELDKWSLWNNFNI 52
Db 420 NHTTWLEWREINNYTSLIHSLEESQKQKNEQLLELDKWSLWNNFNI 471

RESULT 15

AAW43066
ID AAW43066 standard; peptide; 853 AA.

XX AAW43066;

DT 17-OCT-2003 (revised)

DT 11-SEP-1998 (first entry)

XX HIV-1 gp120 protein fragment from isolate HXB2.

XX gp120 protein; purification; fractionation; ion exchange; chromatography;
KW binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.

XX OS Human immunodeficiency virus 1.

XX US596238-A.

XX 09-DEC-1997.

XX PF 11-MAY-1995; 95US-00439286.

XX PR 20-AUG-1991; 91US-00684963.

XX PR 16-AUG-1993; 93US-00109002.

XX PR 09-MAY-1994; 94US-00240073.

XX PA (CHIR) CHIRON CORP.

XX Scandella C, Haigwood NL;

XX WPI; 1998-041353/04.

XX Purification of HIV gp120 - using chromatographic methods.

XX Disclosure; Fig 2A-W; 53pp; English.

XX AAW43066-W43080 are fragments of the gp120 protein from different human
CC immunodeficiency virus type I (HIV-1) isolates. These proteins are used
CC in a novel method for purifying HIV gp120 so as to provide a purified
CC gp120 glycopeptide having protein/protein binding properties
CC substantially identical to natural viral HIV gp120. The method involves
CC fractionating a crude gp120 preparation containing full-length,
CC glycosylated gp120 using ion exchange chromatography so as to provide a
CC first collection of fractions. A fraction from the first collection is
CC selected that exhibits specific binding affinity for CD4 peptide, thereby
CC producing a first fractionated material. The first fractionated material
CC is fractionated by hydrophobic interaction chromatography so as to
CC provide a second collection of fractions from which a second collection
CC is selected that exhibits specific binding affinity for CD4 peptide. This
CC second fraction is fractionated by size exclusion chromatography so as to
CC provide a third collection of fractions exhibiting specific binding
CC affinity for CD4 peptide, thereby providing the purified gp120. The
CC purified gp120 can be used for antibody production and in vaccines.
CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 853 AA;

Query Match 64.2%; Score 290; DB 2; Length 853;

Best Local Similarity 98.1%; Pred. No. 9.8e-20;

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NHTTWLEWREINNYTSLIHSLEESQKQKNEQLLELDKWSLWNNFNI 52
Db 621 NHTTWLEWREINNYTSLIHSLEESQKQKNEQLLELDKWSLWNNFNI 672

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:37:04 ; Search time 33.8924 Seconds
(without alignments)
129.475 Million cell updates/sec

Title: US-09-877-606-4
Perfect score: 452
Sequence: 1 NHTTWLEWDREINNTSLIH.....SKIYHIENIARIKKLIGER 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 290 | 64.2 | 138 | 4 | US-09-570-921-20 |
| 2 | 290 | 64.2 | 138 | 4 | US-09-570-921-21 |
| 3 | 290 | 64.2 | 268 | 3 | US-08-965-056-16 |
| 4 | 290 | 64.2 | 345 | 3 | US-09-272-342B-7 |
| 5 | 290 | 64.2 | 519 | 1 | US-08-589-446-8 |
| 6 | 290 | 64.2 | 519 | 1 | US-08-444-882-8 |
| 7 | 290 | 64.2 | 519 | 2 | US-08-389-459A-8 |
| 8 | 290 | 64.2 | 519 | 3 | US-08-967-867A-8 |
| 9 | 290 | 64.2 | 856 | 2 | US-07-916-098A-2 |
| 10 | 290 | 64.2 | 856 | 4 | US-09-337-387-11 |
| 11 | 278 | 61.5 | 839 | 3 | US-08-472-240A-10 |
| 12 | 277 | 61.3 | 56 | 4 | US-09-779-451-4 |
| 13 | 277 | 61.3 | 138 | 4 | US-09-570-921-22 |
| 14 | 277 | 61.3 | 138 | 4 | US-09-570-921-24 |
| 15 | 277 | 61.3 | 138 | 4 | US-09-570-921-26 |
| 16 | 277 | 61.3 | 138 | 4 | US-09-570-921-58 |
| 17 | 277 | 61.3 | 237 | 3 | US-08-388-353-641 |
| 18 | 277 | 61.3 | 237 | 3 | US-08-488-551B-641 |
| 19 | 277 | 61.3 | 268 | 3 | US-08-965-056-17 |
| 20 | 277 | 61.3 | 268 | 3 | US-08-965-056-18 |
| 21 | 277 | 61.3 | 282 | 5 | PCT-US95-13335-1 |
| 22 | 277 | 61.3 | 338 | 3 | US-08-486-099-90 |
| 23 | 277 | 61.3 | 338 | 3 | US-08-360-107A-100 |
| 24 | 277 | 61.3 | 338 | 3 | US-08-484-223B-90 |
| 25 | 277 | 61.3 | 338 | 3 | US-08-919-597-90 |
| 26 | 277 | 61.3 | 338 | 3 | US-08-475-668A-90 |
| 27 | 277 | 61.3 | 338 | 3 | US-08-485-551A-90 |

| | | | | | | |
|----|-----|------|-----|---|-------------------|-------------------|
| 28 | 277 | 61.3 | 338 | 3 | US-08-471-913A-90 | Sequence 90, Appl |
| 29 | 277 | 61.3 | 338 | 3 | US-08-485-264A-90 | Sequence 90, Appl |
| 30 | 277 | 61.3 | 338 | 4 | US-08-474-349A-90 | Sequence 90, Appl |
| 31 | 277 | 61.3 | 338 | 4 | US-08-255-208A-26 | Sequence 26, Appl |
| 32 | 277 | 61.3 | 338 | 4 | US-08-470-896-90 | Sequence 90, Appl |
| 33 | 277 | 61.3 | 338 | 4 | US-08-485-546A-90 | Sequence 90, Appl |
| 34 | 277 | 61.3 | 345 | 4 | US-08-817-441-49 | Sequence 49, Appl |
| 35 | 277 | 61.3 | 345 | 4 | US-09-779-451-8 | Sequence 8, Appl |
| 36 | 277 | 61.3 | 615 | 3 | US-09-257-490-11 | Sequence 11, Appl |
| 37 | 277 | 61.3 | 826 | 1 | US-08-375-510-2 | Sequence 2, Appl |
| 38 | 277 | 61.3 | 826 | 2 | US-08-487-657-2 | Sequence 2, Appl |
| 39 | 277 | 61.3 | 854 | 4 | US-09-309-572-23 | Sequence 23, Appl |
| 40 | 277 | 61.3 | 854 | 4 | US-09-718-096-23 | Sequence 23, Appl |
| 41 | 277 | 61.3 | 856 | 3 | US-09-124-900-9 | Sequence 9, Appl |
| 42 | 277 | 61.3 | 861 | 1 | US-08-127-499A-14 | Sequence 14, Appl |
| 43 | 277 | 61.3 | 861 | 1 | US-08-482-847-14 | Sequence 14, Appl |
| 44 | 277 | 61.3 | 861 | 3 | US-07-956-483-10 | Sequence 10, Appl |
| 45 | 277 | 61.3 | 861 | 3 | US-07-956-483-16 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1

US-09-570-921-20
; Sequence 20, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-20

Query Match 64.2%; Score 290; DB 4; Length 138;
Best Local Similarity 98.1%; Pred. No. 2.8e-22;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLIEESQKQKNEQELLELDKWSLWNWFI 52
|||||
DB 85 NHTTWLEWDREINNTSLIHSLIEESQKQKNEQELLELDKWSLWNWFI 136

RESULT 2

US-09-570-921-21
; Sequence 21, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21

```
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-21

Query Match      64.2%; Score 290; DB 4; Length 138;
Best Local Similarity 98.1%; Pred. No. 2.8e-22;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNNFNI 52
    |||||:|||||
Db 85 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNNFNI 136
    |||||:|||||

RESULT 3
US-08-965-056-16
; Sequence 16, Application US/08965056
; Patent No. 6271198
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; APPLICANT: J. Kevin Judice
; APPLICANT: Robert S. McDowell
; APPLICANT: J. Christopher Phelan
; APPLICANT: Melissa A. Starovassnik
; APPLICANT: James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; TITLE OF INVENTION: Making Same
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,056
; FILING DATE: 05-No. 6271198-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-965-056-16

Query Match      64.2%; Score 290; DB 3; Length 268;
Best Local Similarity 98.1%; Pred. No. 5.9e-22;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNNFNI 52
    |||||:|||||
Db 154 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNNFNI 205
    |||||:|||||

RESULT 4
US-09-272-342B-7
; Sequence 7, Application US/09272342B
; Patent No. 6294341
; GENERAL INFORMATION:
; APPLICANT: YU, YEON-GYU

; APPLICANT: KIM, SUNG-HOU
; APPLICANT: RYU, JAE-RYEON
; TITLE OF INVENTION: METHOD FOR DETECTING A SUBSTANCE HAVING AN ACTIVITY TO
; TITLE OF INVENTION: INHIBIT HIV INFECTION USING IMMUNOASSAY AND VARIANT
; FILE REFERENCE: 2901-0125-0
; CURRENT APPLICATION NUMBER: US/09/272,342B
; CURRENT FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-272-342B-7

Query Match      64.2%; Score 290; DB 3; Length 345;
Best Local Similarity 98.1%; Pred. No. 7.9e-22;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNNFNI 52
    |||||:|||||
Db 113 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNNFNI 164
    |||||:|||||

RESULT 5
US-08-589-446-8
; Sequence 8, Application US/08589446
; Patent No. 5614413
; GENERAL INFORMATION:
; APPLICANT: Morrow, Casey D.
; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,446
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,009
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Geary III, William C.
; REGISTRATION NUMBER: 31,359
; REFERENCE/DOCKET NUMBER: URG-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 519 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-589-446-8

Query Match      64.2%; Score 290; DB 1; Length 519;
Best Local Similarity 98.1%; Pred. No. 1.3e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```


QY 1 NHTTWLEWDREINNYTSLIHSLSIESQOQKNEQELLELDKWASLWNWFI 52
 Db 420 NHTTWLEWDREINNYTSLIHSLSIESQOQKNEQELLELDKWASLWNWFI 471

RESULT 6

US-08-444-882-8
 ; Sequence 8, Application US/08444882
 ; Patent No. 5622705
 ; GENERAL INFORMATION:
 ; APPLICANT: Morrow, Casey D.
 ; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
 ; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
 ; TITLE OF INVENTION: USING SAME
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 STATE STREET, SUITE 510
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/444,882
 ; FILING DATE: 19-MAY-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION NUMBER: US 08/087,009
 ; FILING DATE: 01-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Geary III, William C.
 ; REGISTRATION NUMBER: 31,359
 ; REFERENCE/DOCKET NUMBER: UAG-004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 8:
 ; LENGTH: 519 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-444-882-8

Query Match 64.2%; Score 290; DB 1; Length 519;
 Best Local Similarity 98.1%; Pred. No. 1.3e-21;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQOQKNEQELLELDKWASLWNWFI 52
 Db 420 NHTTWLEWDREINNYTSLIHSLSIESQOQKNEQELLELDKWASLWNWFI 471

RESULT 7

US-08-389-459A-8
 ; Sequence 8, Application US/08389459A
 ; Patent No. 5817512
 ; GENERAL INFORMATION:
 ; APPLICANT: Morrow, Casey D. and Porter, Donna, C.
 ; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS
 ; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
 ; TITLE OF INVENTION: USING SAME
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 STATE STREET, SUITE 510
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS

; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/389,459A
 ; FILING DATE: 15-FEB-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER: US 08/087,009
 ; FILING DATE: 01-JUL-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Silveri, Jean M.
 ; REGISTRATION NUMBER: 39,030
 ; REFERENCE/DOCKET NUMBER: UAG-004CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 8:
 ; LENGTH: 519 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-389-459A-8

Query Match 64.2%; Score 290; DB 2; Length 519;
 Best Local Similarity 98.1%; Pred. No. 1.3e-21;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQOQKNEQELLELDKWASLWNWFI 52
 Db 420 NHTTWLEWDREINNYTSLIHSLSIESQOQKNEQELLELDKWASLWNWFI 471

RESULT 8

US-08-987-867A-8
 ; Sequence 8, Application US/08987867A
 ; Patent No. 6063384
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Morrow et al.
 ; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
 ; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
 ; TITLE OF INVENTION: USING SAME
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 STATE STREET
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/987,867A
 ; FILING DATE: 09-DEC-1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION NUMBER: US 08/087,009
 ; FILING DATE: 01-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Louis
 ; REGISTRATION NUMBER: 35,965
 ; REFERENCE/DOCKET NUMBER: UAP-004CPDV
 ; TELECOMMUNICATION INFORMATION:

```
/ TELEPHONE: (617) 227-7400
/ TELEFAX: (617) 742-4214
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 519 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-987-867A-8

Query Match      64.2%; Score 290; DB 3; Length 519;
Best Local Similarity 98.1%; Pred. No. 1.3e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NHTTWLEWDREINNTSLIHSLSIESNQKQKNEQLLELDKWSLWNWNI 52
Db      420 NHTTWLEWDREINNTSLIHSLSIESNQKQKNEQLLELDKWSLWNWNI 471

RESULT 9
US-07-916-098A-2
; Sequence 2, Application US/07916098A
; Patent No. 5871732
; GENERAL INFORMATION:
; APPLICANT: BURKLY, LINDA C.
; APPLICANT: CHISHOLM, PATRICIA L.
; APPLICANT: THOMAS, DAVID W.
; APPLICANT: ROSA, MARGARET D.
; APPLICANT: ROSA, JOSEPH J.
; TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
; TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
; STREET: 10 SOUTH WACKER DRIVE
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/916,098A
; FILING DATE: July 24, 1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08843
; FILING DATE: No. 5871732ember 27, 1991
; CLASSIFICATION: 424
; APPLICATION NUMBER: 07/618,542
; FILING DATE: No. 5871732ember 27, 1990
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: JOHN J. MC DONNELL
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,310-G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 715-1000
; TELEFAX: (312) 715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-916-098A-2

Query Match      64.2%; Score 290; DB 2; Length 856;
Best Local Similarity 98.1%; Pred. No. 2.3e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NHTTWLEWDREINNTSLIHSLSIESNQKQKNEQLLELDKWSLWNWNI 52
Db      624 NHTTWLEWDREINNTSLIHSLSIESNQKQKNEQLLELDKWSLWNWNI 675

RESULT 10
US-09-337-387-11
; Sequence 11, Application US/09337387
; Patent No. 6420545
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/09/337,387
; CURRENT FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-337-387-11

Query Match      64.2%; Score 290; DB 4; Length 856;
Best Local Similarity 98.1%; Pred. No. 2.3e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NHTTWLEWDREINNTSLIHSLSIESNQKQKNEQLLELDKWSLWNWNI 52
Db      624 NHTTWLEWDREINNTSLIHSLSIESNQKQKNEQLLELDKWSLWNWNI 675

RESULT 11
US-08-472-240A-10
; Sequence 10, Application US/08472240A
; Patent No. 6284248
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: gp160 VARIANT
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,240A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,483
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-055
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 839 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-08-472-240A-10

Query Match          61.3%; Score 278; DB 3; Length 839;
Best Local Similarity 68.0%; Pred. No. 3.6e-20;
Matches 51; Conservative 8; Mismatches 4; Indels 12; Gaps 1;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 60
Db 629 NNTWMEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFN----- 679

QY 61 IERILSKIYHENEI 75
Db 680 ---ITNWLWYIKNRV 691

RESULT 12
US-09-779-451-4
; Sequence 4, Application US/09779451
; Patent No. 6605427
; GENERAL INFORMATION:
; APPLICANT: Wild, Carl T.
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors
; FILE REFERENCE: 1900.0300003
; CURRENT APPLICATION NUMBER: US/09/779,451
; CURRENT FILING DATE: 2001-08-17
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/235,901
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/181,543
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-779-451-4

Query Match          61.3%; Score 277; DB 4; Length 56;
Best Local Similarity 94.2%; Pred. No. 2e-21;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 52
Db 2 NNTWMEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 53

RESULT 13
US-09-570-921-22
; Sequence 22, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
;   TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22

Query Match          61.3%; Score 277; DB 4; Length 138;
Best Local Similarity 94.2%; Pred. No. 5.7e-21;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-22

Query Match          61.3%; Score 277; DB 4; Length 138;
Best Local Similarity 94.2%; Pred. No. 5.7e-21;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 52
Db 85 NNTWMEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 136

RESULT 14
US-09-570-921-24
; Sequence 24, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
;   TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-24

Query Match          61.3%; Score 277; DB 4; Length 138;
Best Local Similarity 94.2%; Pred. No. 5.7e-21;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 52
Db 85 NNTWMEWDREINNYTSLIHSLSIESQNOQEKNEQELLELDKWASLWNWFNI 136

RESULT 15
US-09-570-921-26
; Sequence 26, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
;   TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; CURRENT FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-26

Query Match          61.3%; Score 277; DB 4; Length 138;
Best Local Similarity 94.2%; Pred. No. 5.7e-21;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 NHTTWLEWDREINNYTSLIHSLIEESQOQEKNEQELLELDKWASLWNWFNI 52
| : |||||
Db 85 NNTWMEWDREINNYTSLIHSLIEESQOQEKNEQELLELDKWASLWNWFNI 136

Search completed: February 25, 2004, 15:46:54
Job time : 33.8924 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:44:50 ; Search time 71.0127 Seconds
(without alignments)
252.744 Million cell updates/sec

Title: US-09-877-606-4

Perfect score: 452

Sequence: 1 NHTTWLEWDREINNTSLIH.....SKIYHIENRIARKKLIGR 85

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 290 | 64.2 | 268 | 9 | US-09-854-816-16 |
| 2 | 290 | 64.2 | 344 | 14 | US-10-040-349B-1 |
| 3 | 290 | 64.2 | 359 | 14 | US-10-214-670-58 |
| 4 | 290 | 64.2 | 519 | 9 | US-09-756-551A-8 |
| 5 | 290 | 64.2 | 856 | 9 | US-09-476-242-1 |
| 6 | 290 | 64.2 | 856 | 14 | US-10-196-515-11 |
| 7 | 278 | 61.5 | 177 | 14 | US-10-040-349B-2 |
| 8 | 278 | 61.5 | 200 | 14 | US-10-263-103-25 |
| 9 | 278 | 61.5 | 200 | 15 | US-10-438-691-8 |
| 10 | 277 | 61.3 | 56 | 9 | US-09-779-451-4 |
| 11 | 277 | 61.3 | 221 | 14 | US-10-059-271-84 |
| 12 | 277 | 61.3 | 232 | 14 | US-10-059-271-81 |
| 13 | 277 | 61.3 | 254 | 14 | US-10-059-271-82 |
| 14 | 277 | 61.3 | 256 | 14 | US-10-059-271-97 |
| 15 | 277 | 61.3 | 268 | 9 | US-09-854-816-17 |

| | | | | | |
|----|-------|------|------|----|-------------------|
| 16 | 277 | 61.3 | 268 | 9 | US-09-854-816-18 |
| 17 | 277 | 61.3 | 345 | 9 | US-09-779-451-8 |
| 18 | 277 | 61.3 | 345 | 14 | US-10-026-741-49 |
| 19 | 277 | 61.3 | 391 | 14 | US-10-059-271-93 |
| 20 | 277 | 61.3 | 853 | 13 | US-10-003-035-33 |
| 21 | 277 | 61.3 | 853 | 14 | US-10-286-332A-33 |
| 22 | 277 | 61.3 | 853 | 15 | US-10-280-915-33 |
| 23 | 277 | 61.3 | 861 | 14 | US-10-026-741-103 |
| 24 | 277 | 61.3 | 1101 | 13 | US-10-003-035-53 |
| 25 | 277 | 61.3 | 1101 | 14 | US-10-286-332A-53 |
| 26 | 277 | 61.3 | 1101 | 15 | US-10-280-915-53 |
| 27 | 277 | 61.3 | 1186 | 13 | US-10-003-035-55 |
| 28 | 277 | 61.3 | 1186 | 14 | US-10-286-332A-55 |
| 29 | 277 | 61.3 | 1186 | 15 | US-10-280-915-55 |
| 30 | 274 | 60.6 | 268 | 9 | US-09-854-816-19 |
| 31 | 267 | 59.1 | 58 | 14 | US-10-252-136-11 |
| 32 | 266 | 58.8 | 759 | 14 | US-10-196-515-12 |
| 33 | 262 | 58.0 | 726 | 14 | US-10-196-515-3 |
| 34 | 261 | 57.7 | 48 | 14 | US-10-351-641-547 |
| 35 | 260 | 57.5 | 268 | 9 | US-09-854-816-13 |
| 36 | 260 | 57.5 | 269 | 9 | US-09-854-816-46 |
| 37 | 257.5 | 57.0 | 108 | 14 | US-10-263-103-33 |
| 38 | 256 | 56.6 | 46 | 9 | US-09-779-451-41 |
| 39 | 255 | 56.4 | 269 | 9 | US-09-854-816-43 |
| 40 | 254 | 56.2 | 268 | 9 | US-09-854-816-9 |
| 41 | 253 | 56.0 | 269 | 9 | US-09-854-816-12 |
| 42 | 252 | 55.8 | 268 | 9 | US-09-854-816-26 |
| 43 | 252 | 55.8 | 269 | 9 | US-09-854-816-28 |
| 44 | 252 | 55.8 | 842 | 14 | US-10-190-435-2 |
| 45 | 252 | 55.8 | 842 | 14 | US-10-241-009-2 |

ALIGNMENTS

RESULT 1

US-09-854-816-16
; Sequence 16, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovashnik
; James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Phd., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2-
; TELECOMMUNICATION INFORMATION:

```

;
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-854-816-16

Query Match 64.2%; Score 290; DB 9; Length 268;
Best Local Similarity 98.1%; Pred. No. 4.9e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 52
Db 154 NHTTWMEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 205

RESULT 2
US-10-040-349B-1
; Sequence 1, Application US/10040349B
; Publication No. US20030082521A1
; GENERAL INFORMATION:
; APPLICANT: Brasseur, Robert
; APPLICANT: Charlotaux, Benoit
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; TITLE OF INVENTION: Polypeptide Inducing Antibodies Neutralizing HIV
; FILE REFERENCE: 01-078-A
; CURRENT APPLICATION NUMBER: US/10/040,349B
; CURRENT FILING DATE: 2002-07-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: (1)..(344)
; OTHER INFORMATION: gp41 LAI protein
US-10-040-349B-1

Query Match 64.2%; Score 290; DB 14; Length 344;
Best Local Similarity 98.1%; Pred. No. 6.5e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 52
Db 113 NHTTWMEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 164

RESULT 3
US-10-214-670-58
; Sequence 58, Application US/10214670
; Publication No. US20030180715A1
; GENERAL INFORMATION:
; APPLICANT: Tibotec Pharmaceuticals Ltd.
; TITLE OF INVENTION: Methods and means for assessing HIV envelope inhibitor
; FILE REFERENCE: VIP-0021 seq listing
; CURRENT APPLICATION NUMBER: US/10/214,670
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: EP 01203011.0
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/310497
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 359

;
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-214-670-58

Query Match 64.2%; Score 290; DB 14; Length 359;
Best Local Similarity 98.1%; Pred. No. 6.8e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 52
Db 135 NHTTWMEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 186

RESULT 4
US-09-756-551A-8
; Sequence 8, Application US/09756551A
; Patent No. US20020051768A1
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSER: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
; FILING DATE: 08-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/376,184
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 519 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-756-551A-8

Query Match 64.2%; Score 290; DB 9; Length 519;
Best Local Similarity 98.1%; Pred. No. 1e-20;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQEKNEQLLELDKQWASLWNNWFI 52
```

Db 420 NHHTWWDREINNYTSLIHSLIEESONQQRNEQELLELDKWSLNWFNI 471

RESULT 5

```

US-09-476-242-1
; Sequence No. Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Eric
; APPLICANT: MARTIN, Karin
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476.242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 1
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-1

```

| | | | | |
|--------------------------|-------|--------------------|-----------|-------------|
| Query Match | 64.2% | Score 290; | DB 9; | Length 856; |
| Best Local Similarity | 98.1% | Pred. No. 1.8e-20; | | |
| Matches 51; Conservative | | 1; Mismatches 0; | Indels 0; | Gaps 0; |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|-----|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|
| Qy | 1 | NH | T | W | L | E | W | D | R | E | I | N | N | Y | T | S | L | I | E | S | Q | N | Q | E | K | N | E | Q | E | L | L | E | D | K | W | A | S | L | N | W | F | N | I | 52 |
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | | | | |
| Db | 624 | NH | T | W | L | E | W | D | R | E | I | N | N | Y | T | S | L | I | E | S | Q | N | Q | E | K | N | E | Q | E | L | L | E | D | K | W | A | S | L | N | W | F | N | I | 675 |

RESULT 6

```

US-10-196-515-11
/ Sequence 11, Application US/10196515
/ Publication No. US20030091594A1
/ GENERAL INFORMATION:
/ APPLICANT: HOXIE, James A.
/ APPLICANT: LABRANCHE, Celia C.
/ APPLICANT: DOMS, Robert W. L.
/ APPLICANT: HOFFMAN, Trevor L.
/ TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
/ TITLE OF INVENTION: THERAPEUTICS
/ FILE REFERENCE: Hoxie 9596-104U1 (0282)
/ CURRENT APPLICATION NUMBER: US/10/196,515
/ CURRENT FILING DATE: 2002-07-16
/ PRIOR APPLICATION NUMBER: US/09/337,387
/ PRIOR FILING DATE: 1999-06-22
/ PRIOR APPLICATION NUMBER: US 09/317,556
/ PRIOR FILING DATE: 1999-05-24
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 11
/ LENGTH: 856
/ TYPE: PRT
/ ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

```

| | | | | | | | |
|-----------------------|-------|--------------|----------|------------|-----|--------|------|
| Query Match | 64.2% | Score | 280; | DB | 14; | Length | 856; |
| Best Local Similarity | 98.1% | Pred. No. | 1.8e-20; | | | | |
| Matches | 51; | Conservative | 1; | Mismatches | 0; | Indels | 0; |
| | | | | | | Gaps | 0; |

QY 1 NHTTWLEWDREINNYTSLIHSLSIESQNOEKNEQELLELDKWSLWNWFI 52
 |||||:|||||
 Db 624 NHTTWLEWDREINNYTSLIHSLSIESQNOEKNEQELLELDKWSLWNWFI 675

RESIT.T 7

US-10-040-349B-2
; Sequence 2, Application US/10040349B
; Publication No. US2003008252A1
; GENERAL INFORMATION:

```

; APPLICANT: Brasseur, Robert
; APPLICANT: Charlotiaux, Benoit
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; TITLE OF INVENTION: Polypeptide Inducing Antibodies Neutralizing HIV
; FILE REFERENCES: 01-078-A
; CURRENT APPLICATION NUMBER: US/10/040,349B
; CURRENT FILING DATE: 2002-07-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 177
; TYPE: PRP
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: (1)..(177)
; OTHER INFORMATION: polypeptide derived from gp41 LAI
US-10-040-349B-2

```

```
Query Match      61.5%; Score 278; DB 14; Length 177;
Best Local Similarity 68.0%; Pred. No. 4.8e-20;
Matches 51; Conservative 8; Mismatches 4; Indels 12; Gaps 1;

Qy    1 NHTTWLEWDREINNYTSLIHSLEEESNQOQEKNQEQLLELDKWLASLWNNWFNIKIQIEDK 60
      | : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db    90 NNMTWMEWDREINNYTSLIHSLEEESNQOQEKNQEQLLELDKWLASLWNNWFN----- 140

Qy    61 IEEILSKIYHIENEI 75
      | : ::|| : :
Db    141 ---ITNWLYIKNRV 152
```

RESULT 8

```

US-10-263-103-25
; Sequence 25, Application US/10263103
; Publication No. US20030138445A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PASTEUR
; APPLICANT: Chevalier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; APPLICANT: Sodoyer, Regis
; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: polypeptide
US-10-263-103-25

```

```
Query Match      61.5%; Score 278; DB 14; Length 200;  
Best Local Similarity 68.0%; Pred. No. 5.5e-20;  
Matches 51; Conservative 8; Mismatches 4; Indels 12; Gaps 1;  
  
Qy    1 NHTTWLEWDREINNYTSLIHLSLIEESONQEKNQEQLLEDKVASLNWNFNIIKKIQIEDK 60  
       | :|::|||:|||||:|||||:|||||:|||||:|||||  
Db   113 NNMTWMDWEDEIENYTSLIHLSLIEESONQEKNQEQLLEDKVASLNWNFN----- 163  
  
Qy    61 IEEILSKIYHIENEI 75  
       | : :: || :  
Db   164 ---ITNMLWTYIKRV 175
```

RESULT 9

US-10-438-691-8

```
; Sequence 8, Application US/10438691
; Publication No. US20040009188A1
; GENERAL INFORMATION:
; APPLICANT: REPKE, HEINRICH
; APPLICANT: BUDDÉ, ECKHARD
; APPLICANT: NICOLAUS, STEFAN
; TITLE OF INVENTION: PROTEIN HAVING MULTIPLE ANTIGEN/EPITOPE SEQUENCES AND
; TITLE OF INVENTION: BEING IMMOBILIZED
; FILE REFERENCE: ALBRE-22
; CURRENT APPLICATION NUMBER: US/10/059,271
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: DE 101 06 295
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-059-271-84

Query Match 61.3%; Score 277; DB 15; Length 200;
Best Local Similarity 68.0%; Pred. No. 5.5e-20;
Matches 51; Conservative 8; Mismatches 4; Indels 12; Gaps 1;

QY 1 NHTTWLEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 60
Db 113 NNTWMEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 163

QY 61 IBEILSKIYHIENEI 75
Db 164 ---ITNWLWYIKRV 175

RESULT 10
US-09-779-451-4
; Sequence 4, Application US/09779451
; Patent No. US20020094521A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Carl T.
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors
; FILE REFERENCE: 1900.0300003
; CURRENT APPLICATION NUMBER: US/09/779,451
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/235,901
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/181,543
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-779-451-4

Query Match 61.3%; Score 277; DB 9; Length 56;
Best Local Similarity 94.2%; Pred. No. 1.7e-20;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 52
Db 2 NNTWMEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 53

RESULT 11
US-10-059-271-84
; Sequence 84, Application US/10059271
; Publication No. US20030082208A1
; GENERAL INFORMATION:
; APPLICANT: REPKE, HEINRICH
; APPLICANT: BUDDÉ, ECKHARD
```

```
; GENERAL INFORMATION:
; APPLICANT: REPKE, HEINRICH
; APPLICANT: BUDDÉ, ECKHARD
; APPLICANT: NICOLAUS, STEFAN
; TITLE OF INVENTION: PROTEIN HAVING MULTIPLE ANTIGEN/EPITOPE SEQUENCES AND
; TITLE OF INVENTION: BEING IMMOBILIZED
; FILE REFERENCE: ALBRE-22
; CURRENT APPLICATION NUMBER: US/10/059,271
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: DE 101 06 295
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-059-271-84

Query Match 61.3%; Score 277; DB 14; Length 221;
Best Local Similarity 94.2%; Pred. No. 7.7e-20;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 52
Db 117 NNTWMEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 168

RESULT 12
US-10-059-271-81
; Sequence 81, Application US/10059271
; Publication No. US20030082208A1
; GENERAL INFORMATION:
; APPLICANT: REPKE, HEINRICH
; APPLICANT: BUDDÉ, ECKHARD
; APPLICANT: NICOLAUS, STEFAN
; TITLE OF INVENTION: PROTEIN HAVING MULTIPLE ANTIGEN/EPITOPE SEQUENCES AND
; TITLE OF INVENTION: BEING IMMOBILIZED
; FILE REFERENCE: ALBRE-22
; CURRENT APPLICATION NUMBER: US/10/059,271
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: DE 101 06 295
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-059-271-81

Query Match 61.3%; Score 277; DB 14; Length 232;
Best Local Similarity 94.2%; Pred. No. 8.2e-20;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 52
Db 130 NNTWMEWDREINNTSLIHSLEESQKQKNEQELLELDKWSLWNNWFI 181

RESULT 13
US-10-059-271-82
; Sequence 82, Application US/10059271
; Publication No. US20030082208A1
; GENERAL INFORMATION:
; APPLICANT: REPKE, HEINRICH
; APPLICANT: BUDDÉ, ECKHARD
```



```

; APPLICANT: NICOLAUS, STEFAN
; TITLE OF INVENTION: PROTEIN HAVING MULTIPLE ANTIGEN/EPITOPE SEQUENCES AND
; FILE REFERENCE: ALBRE-22
; CURRENT APPLICATION NUMBER: US/10/059,271
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: DE 101 06 295
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-059-271-82

Query Match 61.3%; Score 277; DB 14; Length 254;
Best Local Similarity 94.2%; Pred. No. 9e-20;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFNI 52
Db 152 NNTWTWEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFNI 203

RESULT 14
US-10-059-271-97
; Sequence 97, Application US/10059271
; Publication No.: US20030082208A1
; GENERAL INFORMATION:
; APPLICANT: REPKE, HEINRICH
; APPLICANT: BUDDÉ, ECKHARD
; APPLICANT: NICOLAUS, STEFAN
; TITLE OF INVENTION: PROTEIN HAVING MULTIPLE ANTIGEN/EPITOPE SEQUENCES AND
; FILE REFERENCE: ALBRE-22
; CURRENT APPLICATION NUMBER: US/10/059,271
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: DE 101 06 295
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-059-271-97

Query Match 61.3%; Score 277; DB 14; Length 256;
Best Local Similarity 94.2%; Pred. No. 9.1e-20;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFNI 52
Db 152 NNTWTWEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFNI 203

RESULT 15
US-09-854-816-17
; Sequence 17, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovasnik
```

```

; James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; Making Same
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-854-816-17

Query Match 61.3%; Score 277; DB 9; Length 268;
Best Local Similarity 94.2%; Pred. No. 9.6e-20;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFNI 52
Db 154 NNTWTWEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFNI 205

Search completed: February 25, 2004, 16:04:25
Job time : 72.0127 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:36:14 ; Search time 27.4367 Seconds
(without alignments)
298.005 Million cell updates/sec

Title: US-09-877-606-4
Perfect score: 452
Sequence: 1 NHTTWLEWDREINNYTSLIH.....SKYHIENIARIKKLIGER 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|------------------------|
| 1 | 277 | 61.3 | 851 | 2 S33985 | env polyprotein - |
| 2 | 277 | 61.3 | 854 | 2 S13288 | env protein - huma |
| 3 | 277 | 61.3 | 856 | 1 VCLJH3 | env polyprotein pr |
| 4 | 277 | 61.3 | 861 | 1 VCLJLV | env polyprotein pr |
| 5 | 268 | 59.3 | 856 | 1 VCLJVL | env polyprotein pr |
| 6 | 261 | 57.7 | 358 | 2 S21998 | env polyprotein g |
| 7 | 254 | 56.2 | 856 | 1 VCLJ3W | env polyprotein pr |
| 8 | 253 | 56.0 | 861 | 1 VCLJSC | env polyprotein pr |
| 9 | 251 | 55.5 | 443 | 2 C4121 | env polyprotein P |
| 10 | 247 | 54.6 | 847 | 2 T09448 | env polyprotein P |
| 11 | 247 | 54.6 | 847 | 2 S13289 | env polyprotein - huma |
| 12 | 246 | 54.4 | 852 | 2 T12016 | env polyprotein g |
| 13 | 245 | 54.2 | 357 | 2 S21996 | env polyprotein g |
| 14 | 244 | 54.0 | 358 | 2 S22002 | env polyprotein g |
| 15 | 244 | 54.0 | 358 | 2 S22000 | env polyprotein g |
| 16 | 244 | 54.0 | 358 | 2 S70417 | env polyprotein g |
| 17 | 243 | 53.8 | 859 | 1 VCLJMN | env polyprotein pr |
| 18 | 242 | 53.5 | 357 | 2 S22006 | env polyprotein g |
| 19 | 241 | 53.3 | 357 | 2 S21994 | env polyprotein g |
| 20 | 241 | 53.3 | 853 | 2 S54384 | env polyprotein g |
| 21 | 241 | 53.3 | 855 | 1 VCLJZR | env polyprotein pr |
| 22 | 240 | 53.1 | 357 | 2 S21992 | env polyprotein pr |
| 23 | 239 | 52.9 | 852 | 1 VCLJBR | env polyprotein g |
| 24 | 237 | 52.4 | 357 | 2 S22004 | env polyprotein g |
| 25 | 237 | 52.4 | 445 | 2 A41621 | env polyprotein M |
| 26 | 236 | 52.2 | 729 | 1 VCLJXK | env polyprotein pr |
| 27 | 236 | 52.2 | 855 | 1 VCLJAJ | env polyprotein pr |
| 28 | 236 | 52.2 | 861 | 1 VCLJKB | env polyprotein pr |
| 29 | 236 | 52.2 | 868 | 1 VCLJHA | env polyprotein - |

| | | | | | |
|----|-------|------|-----|----------|--------------------|
| 30 | 234 | 51.8 | 846 | 1 VCLJND | env polyprotein pr |
| 31 | 233 | 51.5 | 843 | 1 H44001 | env polyprotein pr |
| 32 | 230 | 50.9 | 859 | 2 T01672 | env polyprotein pr |
| 33 | 228 | 50.4 | 136 | 2 JU0266 | env polyprotein pr |
| 34 | 228 | 50.4 | 454 | 2 B41621 | env polyprotein D |
| 35 | 223 | 49.3 | 136 | 2 JT0954 | env polyprotein pr |
| 36 | 215 | 47.6 | 856 | 1 A44963 | env polyprotein pr |
| 37 | 213 | 47.1 | 357 | 2 S21990 | env polyprotein g |
| 38 | 198 | 43.8 | 854 | 1 VCLJSI | env polyprotein pr |
| 39 | 197 | 43.6 | 877 | 2 S49197 | env polyprotein p |
| 40 | 181 | 40.0 | 863 | 2 A53034 | gag polyprotein - |
| 41 | 137.5 | 30.4 | 881 | 1 VCLJG3 | env polyprotein - |
| 42 | 136 | 30.1 | 281 | 1 RGBYA2 | amino acid biosynt |
| 43 | 133.5 | 29.5 | 151 | 2 S30448 | env protein - huma |
| 44 | 133.5 | 29.5 | 151 | 2 S30452 | env protein - huma |
| 45 | 133.5 | 29.5 | 881 | 2 S03068 | env protein - huma |

ALIGNMENTS

RESULT 1

S33985
env polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C;Accession: S33985

R;Carlini, F.
submitted to the EMBL Data Library, November 1991

A;Reference number: S33979

A;Accession: S33985

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-851 <CAR>

A;Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77628.1; PID:g60199

C;Superfamily: type B retrovirus env polyprotein

Query Match 61.3%; Score 277; DB 2; Length 851;
Best Local Similarity 94.2%; Pred. No. 4.4e-16;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLIEESQNOQEKNEQLLELDKWASLWNWNI 52
Db 619 NNTWMEWDREINNYTSLIHSLIEESQNOQEKNEQLLELDKWASLWNWNI 670

RESULT 2

S13288
env protein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S13288

R;O'Brien, W.A.; Koyanagi, Y.; Nemazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.

Nature 348, 69-73, 1990

A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120

A;Reference number: S13288; MUID:91043044; PMID:2172833

A;Accession: S13288

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-854 <OBR>

C;Superfamily: type B retrovirus env polyprotein

Query Match 61.3%; Score 277; DB 2; Length 854;
Best Local Similarity 94.2%; Pred. No. 4.5e-16;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLIEESQNOQEKNEQLLELDKWASLWNWNI 52
Db 622 NNTWMEWDREINNYTSLIHSLIEESQNOQEKNEQLLELDKWASLWNWNI 673

RESULT 3

VCLJH3

[illegible]

A;Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity
A;Reference number: A41621; MUID:92107924; PMID:1763038
A;Accession: C41621
A:Molecule type: DNA
A;Residues: 1-443 <BUR>
A;Cross-references: GB:M77230; NID:g328631; PIDN:AAB03792.1; PID:g555015
A;Note: this virus was isolated from the mother's sexual partner
C;Genetics:
A;Gene: env
C;Superfamily: type B retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp.
F;1-25/Product: coat protein gp120 (fragment) #status predicted <GP1>
F;252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>
P;424-443/Domain: transmembrane #status predicted <TMW>
P;9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site: c.

Query Match 55.5%; Score 251; DB 2; Length 443;
Best Local Similarity 57.8%; Pred. No. 3.5e-14;
Matches 49; Conservative 13; Mismatches 11; Indels 12; Gaps 1;

QY 1 NHTTWLEWDREINNYTSLHLSIESQNQEKNQELLELDKWASLWNWFNIKIQIEDK 60
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 364 NNNTWWQWEREINYTSLYNLIESQNQEKNQELLELDKWASLWNWFNS----- 414

QY 61 IBEILSKIHYIENIEARIKKLIGER 85
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 415 ---ITNWMLWKIFIIVGGLIGLR 436

RESULT 10
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C;Accession: T09448
R;Bang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
submitted to the EMBL Data Library, July 1996
A;Reference number: Z16673
A;Accession: T09448
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-847 <PAN>
A;Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781
C;Genetics:
A;Gene: env
C;Superfamily: type B retrovirus env polyprotein

Query Match 54.6%; Score 247; DB 2; Length 847;
Best Local Similarity 82.7%; Pred. No. 1.6e-13;
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLHLSIESQNQEKNQELLELDKWASLWNWFNI 52
 |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 615 NNMTWMEWRBIDNYTSEITYTLIESQNQEKNQELLELDKWASLWNWFPI 666

RESULT 11
S13289
env protein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S13289
R;O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.
Nature 348, 69-73, 1990
A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A;Reference number: S13288; MUID:91043044; PMID:2172833
A;Accession: S13289
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-847 <OB>
C;Superfamily: type B retrovirus env polyprotein

Query Match 54.6%; Score 247; DB 2; Length 847;

```
Best Local Similarity 82.7%; Pred. No. 1.6e-13;
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSIIIESQOQKNEQELLELDKWASLWNWFI 52
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 615 NNTWMEWEREDINTYTSIYTLIESQOQKNEQELLELDKWASLWNWFI 666

RESULT 12
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g2351784
C:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 54.4%; Score 246; DB 2; Length 852;
Best Local Similarity 57.6%; Pred. No. 2e-13;
Matches 49; Conservative 12; Mismatches 12; Indels 12; Gaps 1;

QY 1 NHTTWLEWDREINNTSLIHSIIIESQOQKNEQELLELDKWASLWNWFIKIKQIEDK 60
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 620 NNTWMEWEREDINTYTSIYTLIESQOQKNEQELLELDKWASLWWSF----- 669

QY 61 IBEILSKIYHIENEIARIKKLIGER 85
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 670 --DITKWLWYIKFIMVGLGLR 692

RESULT 13
S21996
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S70422; S21996
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cereb
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <ST2>
A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43624.1; PID:g1067129
A:Experimental source: patient 27L
A>Note: submitted to the EMBL Data Library, July 1991
C:Superfamily: type E retrovirus env polyprotein

Query Match 54.2%; Score 245; DB 2; Length 357;
Best Local Similarity 56.5%; Pred. No. 8.9e-14;
Matches 48; Conservative 13; Mismatches 12; Indels 12; Gaps 1;

QY 1 NHTTWLEWDREINNTSLIHSIIIESQOQKNEQELLELDKWASLWNWFIKIKQIEDK 60
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 125 DNMTWMEWEREINTYTSIYTLIESQOQKNEQELLELDKWASLWWSF----- 175

QY 61 IBEILSKIYHIENEIARIKKLIGER 85
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 176 ---ITNWLWYIKFIMVGLGLR 197

RESULT 14
```

```
S22002
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 3L
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S22002; S70418
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determ
A:Reference number: S21990
A:Accession: S22002
A:Molecule type: DNA
A:Residues: 1-358 <ST1>
A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cereb
A:Reference number: S70417; MUID:92144209; PMID:1736940
A:Accession: S70418
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333, 'X', 335-358 <ST2>
A:Cross-references: EMBL:X61352; NID:g60186
C:Superfamily: type E retrovirus env polyprotein

Query Match 54.0%; Score 244; DB 2; Length 358;
Best Local Similarity 80.8%; Pred. No. 1.1e-13;
Matches 42; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSIIIESQOQKNEQELLELDKWASLWNWFI 52
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 126 DNMTWMEWDREIHNNTSLIYTLIESQOQKNEQELLELDKWDSLWNWFSI 177

RESULT 15
S22000
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 12-Apr-1995
C:Accession: S22000
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determ
A:Reference number: S21990
A:Accession: S22000
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STE>
A:Cross-references: EMBL:X61351
C:Superfamily: type E retrovirus env polyprotein

Query Match 54.0%; Score 244; DB 2; Length 358;
Best Local Similarity 80.8%; Pred. No. 1.1e-13;
Matches 42; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSIIIESQOQKNEQELLELDKWASLWNWFI 52
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 126 DNMTWMEWDREIHNNTSLIYTLIESQOQKNEQELLELDKWDSLWNWFSI 177

Search completed: February 25, 2004, 15:45:37
Job time : 27.4367 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:39 ; Search time 17.2152 Seconds
(without alignments)
257.096 Million cell updates/sec

Title: US-09-877-606-4
Perfect score: 452
Sequence: 1 NHTTWLEWDREINNYTSLIH.....SKYHIENIEARIKKLIGER 85

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 290 | 64.2 | 856 | 1 ENV_HV1H2 | P04578 human immu |
| 2 | 290 | 64.2 | 856 | 1 ENV_HV1H3 | P04624 human immu |
| 3 | 290 | 64.2 | 856 | 1 ENV_HV1LW | Q70626 human immu |
| 4 | 277 | 61.3 | 851 | 1 ENV_HV1B8 | P04582 human immu |
| 5 | 277 | 61.3 | 856 | 1 ENV_HV1B1 | P03375 human immu |
| 6 | 277 | 61.3 | 861 | 1 ENV_HV1BR | P03377 human immu |
| 7 | 274 | 60.6 | 853 | 1 ENV_HV1MF | P19551 human immu |
| 8 | 268 | 59.3 | 856 | 1 ENV_HV1PV | P03376 human immu |
| 9 | 254 | 56.2 | 856 | 1 ENV_HV1W1 | P31872 human immu |
| 10 | 253 | 56.0 | 856 | 1 ENV_HV1SC | P05877 human immu |
| 11 | 252 | 55.8 | 847 | 1 ENV_HV1S1 | P05878 human immu |
| 12 | 252 | 55.8 | 852 | 1 ENV_HV1S3 | P19549 human immu |
| 13 | 245 | 54.2 | 847 | 1 ENV_HV1W2 | P05880 human immu |
| 14 | 243 | 53.8 | 856 | 1 ENV_HV1M2 | P05877 human immu |
| 15 | 241 | 53.3 | 853 | 1 ENV_HV1Z2 | P12487 human immu |
| 16 | 241 | 53.3 | 855 | 1 ENV_HV1Z6 | P04580 human immu |
| 17 | 241 | 53.3 | 867 | 1 ENV_HV1J3 | P12489 human immu |
| 18 | 240 | 53.1 | 865 | 1 ENV_HV1RH | P04579 human immu |
| 19 | 239 | 52.9 | 852 | 1 ENV_HV1BN | P12488 human immu |
| 20 | 238 | 52.7 | 853 | 1 ENV_HV1EL | P04581 human immu |
| 21 | 236 | 52.2 | 855 | 1 ENV_HV1A2 | P03378 human immu |
| 22 | 236 | 52.2 | 861 | 1 ENV_HV1KB | P31819 human immu |
| 23 | 236 | 52.2 | 868 | 1 ENV_HV1C4 | P05879 human immu |
| 24 | 234 | 51.8 | 846 | 1 ENV_HV1ND | P18799 human immu |
| 25 | 233 | 51.5 | 843 | 1 ENV_HV1Y2 | P35961 human immu |
| 26 | 232 | 51.3 | 848 | 1 ENV_HV1JR | P20871 human immu |
| 27 | 232 | 51.3 | 855 | 1 ENV_HV1OY | P20888 human immu |
| 28 | 230 | 50.9 | 859 | 1 ENV_HV1MA | P04583 human immu |
| 29 | 230 | 50.9 | 863 | 1 ENV_HV1Z8 | P05882 human immu |
| 30 | 215 | 47.6 | 856 | 1 ENV_HV1ZH | P05881 human immu |
| 31 | 198 | 43.8 | 854 | 1 ENV_SIVCZ | P17281 chimpanzee |
| 32 | 142 | 31.4 | 854 | 1 ENV_SIVAI | Q02837 simian immu |
| 33 | 137.5 | 30.4 | 882 | 1 ENV_SIVM1 | P05885 simian immu |

34 136 30.1 281 1 GCN4 YEAST
35 133.5 29.5 881 1 ENV_SIVMK
36 133.5 29.5 885 1 ENV_SIVS4
37 133.5 29.5 889 1 ENV_SIVSP
38 131.5 29.1 712 1 ENV_HV2S2
39 130.5 28.9 859 1 ENV_HV2ST
40 129.5 28.7 859 1 ENV_HV2D2
41 129.5 28.7 880 1 ENV_SIVML
42 128.5 28.4 877 1 ENV_SIVAG
43 124.5 27.5 851 1 ENV_HV2G1
44 124 27.4 768 1 ENV_SIVAI
45 122.5 27.1 851 1 ENV_HV2D1

P03069 saccharomyc
P05884 simian immu
P12492 simian immu
P19503 simian immu
P32536 human immu
P20872 human immu
P15831 human immu
P11267 simian immu
P27977 simian immu
P18040 human immu
P27757 simian immu
P17755 human immu

ALIGNMENTS

RESULT 1
ENV_HV1H2
ID ENV_HV1H2 STANDARD; PRT; 856 AA.
AC P04578; O09779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
virus."
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; K03455; AAB50262.1; -;
DR EMBL; AF038399; AAB99976.1; -;
DR EMBL; AF033819; AAC82596.1; -;
DR PDB; 1DF4; 26-JAN-00.
DR PDB; 1DF5; 26-JAN-00.
DR PDB; 1DLB; 02-APR-00.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.
DR PDB; 1K33; 10-OCT-01.
DR PDB; 1K34; 10-OCT-01.
DR HIV; K03455; ENV8XHB2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure. 1 30
FT SIGNAL 31 511
FT CHAIN 512 856
FT EXTERIOR MEMBRANE GLYCOPROTEIN.
FT TRANSMEMBRANE GLYCOPROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M14100; AAA44679.1; -;
 PDB; 1JAU; 17-OCT-01.
 DR PDB; 1JAV; 17-OCT-01.
 DR HIV; M14100; ENVSHXB3.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 FT SIGNAL 1 30
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 119 205 BY SIMILARITY.
 FT DISULFID 126 196 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 236 331 BY SIMILARITY.
 FT DISULFID 378 445 BY SIMILARITY.
 FT DISULFID 385 418 BY SIMILARITY.
 FT CARBOHYD 88 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 971212 MW; 6FAB16AF85107FE0 CRC64;
 Query Match 64.2%; Score 290; DB 1; Length 856;
 Best Local Similarity 98.1%; Pred. No. 5.6e-18;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NHTTWLEWDREINNTYSLIHSIEESQKQKNEQELLELDKWSLWNWFI 52
 Db 624 NHTTWLEWDREINNTYSLIHSIEESQKQKNEQELLELDKWSLWNWFI 675
 RESULT 2
 ENV HV1H3 STANDARD; PRT; 856 AA.
 AC P04624;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID:11707;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85228248; PubMed=2988795;
 RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
 RA Shaw G.M., Wong-Skaal F., Reddy E.P.;
 RT "HTLV-III env gene products synthesized in E. coli are recognized by
 RL antibodies present in the sera of AIDS patients.";
 RL Cell 41:979-986(1985).
 CC

| ID | ENV_HVILW | STANDARD; | PRT; | 856 AA. |
|----|--|-----------|------|-----------------------------------|
| AC | Q70626; | | | |
| AD | 15-JUL-1998 (Rel. 36, Created) | | | |
| AE | 15-JUL-1998 (Rel. 36, Last sequence update) | | | |
| AF | 10-OCT-2003 (Rel. 42, Last annotation update) | | | |
| AG | Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]. | | | |
| DE | ENV. | | | |
| GN | Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1). | | | |
| OS | Viruses; Retrovirdae; Retroviridae; Lentivirus. | | | |
| OC | NCBI_TaxID=82834; | | | |
| OX | [1] | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RP | MEDLINE=95127297; PubMed=7826699; | | | |
| RX | Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M., Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.; | | | |
| RA | "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HTLV type IIIB)."; | | | |
| RT | AIDS Res. Hum. Retroviruses 10:1143-1155(1994). | | | |
| RL | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; U12055; AAA76690.1; -- | | | |
| DR | PDB; 1IF3; O2-MAY-01. | | | |
| DR | GlycoSuiteDB; Q70626; -- | | | |
| DR | InterPro; IPR000328; Env_GP41. | | | |
| DR | InterPro; IPR000777; GP120. | | | |
| DR | Pfam; PF00516; GP120; 1. | | | |
| DR | Pfam; PF00517; GP41; 1. | | | |
| KW | ADIS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; | | | |
| KW | 3D-structure. | | | |
| FT | SIGNAL | 1 | 30 | |
| FT | CHAIN | 31 | 511 | EXTERIOR MEMBRANE GLYCOPROTEIN. |
| FT | CHAIN | 512 | 856 | TRANSMEMBRANE GLYCOPROTEIN. |
| FT | DT_SULFID | 54 | 74 | BY SIMILARITY. |
| FT | DISULFID | 119 | 205 | BY SIMILARITY. |
| FT | DISULFID | 126 | 196 | BY SIMILARITY. |
| FT | DISULFID | 131 | 157 | BY SIMILARITY. |
| FT | DISULFID | 218 | 247 | BY SIMILARITY. |
| FT | DISULFID | 228 | 239 | BY SIMILARITY. |
| FT | DISULFID | 296 | 331 | BY SIMILARITY. |
| FT | DISULFID | 378 | 445 | BY SIMILARITY. |
| FT | DISULFID | 385 | 418 | BY SIMILARITY. |
| FT | CARBOHYD | 88 | 88 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 136 | 136 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 141 | 141 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 156 | 156 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 160 | 160 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 186 | 186 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 197 | 197 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 230 | 230 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 234 | 234 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 241 | 241 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 262 | 262 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 276 | 276 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 289 | 289 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 295 | 295 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 301 | 301 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 301 | 301 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 332 | 332 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 339 | 339 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 356 | 356 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 386 | 386 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 392 | 392 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 397 | 397 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 406 | 406 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 448 | 448 | N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 463 | 463 | N-LINKED (GLCNAC. .) (POTENTIAL). |

FT DISULFID 378 440 BY SIMILARITY.
 FT DISULFID 385 413 BY SIMILARITY.
 FT CARBOHYD 88 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 61.3%; Score 277; DB 1; Length 851;
 Best Local Similarity 94.2%; Pred. No. 7, 6e-17;
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSIESQKQKNEQKLELDKWSLWNWFI 52
 DB 619 NNMTWLEWDREINNTSLIHSIESQKQKNEQKLELDKWSLWNWFI 670

RESULT 5
 ENV_HV1B1 STANDARD; PRT; 856 AA.
 AC P03375;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 ON NCBI_TaxID=11678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafaleki J.A., Whitehorn E.A., Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L., Lautenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C., Wong-Staal F.;
 RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 RT Nature 313:277-284 (1985).
 RL [2]
 RN
 RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=90285159; PubMed=2355006;
 RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N., Gregory T.J.;
 RA "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in

RT Chinese hamster ovary cells.";
 RL J. Biol. Chem. 265:10373-10382(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M15654; AAA44205.1; .
 DR PIR; A03973; VCLJH3.
 DR HIV; M15654; ENVSBH102.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.
 FT SIGNAL. 1 30
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .)
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .)
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .)
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .)
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .)
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .)
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .)
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .)
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .)
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .)
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .)
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .)
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .)
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .)
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .)
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .)
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .)
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .)
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .)
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .)
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .)
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. .)
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .)
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .)
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97224 MW; 0BFF1A18931BB27 CRC64;
 Query Match 61.3%; Score 277; DB 1; Length 856;
 Best Local Similarity 94.2%; Pred. No. 7, 6e-17;
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSIESQKQKNEQKLELDKWSLWNWFI 52
 DB 624 NNMTWLEWDREINNTSLIHSIESQKQKNEQKLELDKWSLWNWFI 675

```

RESULT 6
ENV HV1BR
ID ENV HV1BR STANDARD; PRT; 861 AA.
AC P0337;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8509333; PubMed=2981635;
RA Main-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL; K02013; BAB59751.1; --
EMBL; A04321; CAA00352.1; --
PIR; A03975; VCLJLV.
DR PDB; 1ERF; 20-FEB-02.
DR HIV; K02013; ENV$BRU.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 516
FT CHAIN 517 861
FT DISULFID 54 74
FT DISULFID 119 210
FT DISULFID 126 201
FT DISULFID 131 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 336
FT DISULFID 383 450
FT DISULFID 390 423
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 235 235
FT CARBOHYD 239 239
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 337 337
FT CARBOHYD 344 344
FT CARBOHYD 361 361
FT CARBOHYD 391 391
FT CARBOHYD 397 397
FT CHAIN 31 516
FT CHAIN 517 861
FT DISULFID 54 74
FT DISULFID 119 210
FT DISULFID 126 201
FT DISULFID 131 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 336
FT DISULFID 383 450
FT DISULFID 390 423
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 235 235
FT CARBOHYD 239 239
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 337 337
FT CARBOHYD 344 344
FT CARBOHYD 361 361
FT CARBOHYD 391 391
FT CARBOHYD 397 397

```

```

FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 642 642 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 861 AA; 97487 MM; 04DE2B4D4E4FD63A CRC64;
Query Match 61.3%; Score 277; DB 1; Length 861;
Best Local Similarity 94.2%; Pred. No. 7,7e-17;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NHTTWLEWDREINNYTSLIHSLEESQNOQKNEQBELLELDKWSLWNNFNI 52
Db 629 NNTWMEWDREINNYTSLIHSLEESQNOQKNEQBELLELDKWSLWNNFNI 680
RESULT 7
ENV HV1MP
ID ENV HV1MP STANDARD; PRT; 853 AA.
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonic C., Mann A.M., Meier C.,
RA Wasiaak A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
RT variants diminished in the ability to induce syncytium-independent
RT cytolysis.";
RL J. Virol. 64:3792-3803(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL; M33943; AAA44850.1; --
DR PDB; 1A1K; 16-JUN-97.
DR HIV; M33943; ENV$MFA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 509
FT CHAIN 510 853
FT DISULFID 54 74
FT DISULFID 119 203
FT DISULFID 126 194
FT DISULFID 131 157
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 329
FT DISULFID 376 443

```

```

FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 814 814 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 33778993B6F22ABA CRC64;

Query Match 60.6%; Score 274; DB 1; Length 853;
Best Local Similarity 92.3%; Pred. No. 1.4e-16;
Matches 48; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 52
Db 622 NMTTWWEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 673

RESULT 8
ENV_HV1PV STANDARD; PRT; 856 AA.
AC F03376;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11700;
RN [1]_TaxID=11700;
RP SEQUENCE FROM N.A.
RX MEDLINE=85111157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A., Capon D.J.;
RT "Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02083; AAB59873.1; -.
DR EMBL; X01762; CAA25903.1; ALT_SEQ.

```

```

DR PIR; A03974; VCLJVL.
DR HIV; K02083; ENVSPV22.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 31 511 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 BY SIMILARITY.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C1209B3 CRC64;

Query Match 59.3%; Score 268; DB 1; Length 856;
Best Local Similarity 90.4%; Pred. No. 4.6e-16;
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 52
Db 624 NMTTWWEWDREINNYTSLIHSLSIESQKQKNEQELLELDKWSLWNWFI 675

RESULT 9
ENV_HV1PV STANDARD; PRT; 856 AA.
AC P31872;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).

```

OC Viruses; Retrovirus; Retroviridae; Lentivirus.
OX NCBI_TaxID=31678;
RN [1]
RP MEDLINE=86218077; PubMed=2423250;
RX Starck B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
RT "Identification and characterization of conserved and variable
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of
RT AIDS";
RL Cell 45:637-648 (1986).
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
DR PIR; A24774; VCUJ3W.
DR PDB; 1LB0; 04-DEC-02.
DR PDB; 1LCX; 04-DEC-02.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 511 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 205 BY SIMILARITY.
FT DISULFID 125 196 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 444 BY SIMILARITY.
FT DISULFID 383 417 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97526 MW; DB68D1E49C404DE9 CRC64;
Query Match 56.2%; Score 254; DB 1; Length 856;
Best Local Similarity 84.6%; Pred. No. 7.7e-15;
Matches 44; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NHTTWLEWDREINNYTSLIHSLEESQKQKNEQELLELDKRWASLNNWFSI 52
Db 624 NNMTWMEWEREIDNYTSLIYNLEESQKQKNEQELLELDKRWASLNNWFSI 675

RESULT 10

ENV_HVISC
ID ENV_HVISC STANDARD; PRT; 856 AA.
AC P05878;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retrovirus; Retroviridae; Lentivirus.
OX NCBI_TaxID=11702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates";
RL Virology 164:531-536(1988).
CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
CC 1984 IN SOUTHERN CALIFORNIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M17450; -; NOT ANNOTATED_CDS.
DR HIV; M17450; ENVSSC.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 29 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 30 510 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 511 856 IN-FRAME TERMINATION CODON.
FT SITE 760 760 BY SIMILARITY.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 160 BY SIMILARITY.
FT DISULFID 219 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 439 BY SIMILARITY.
FT DISULFID 383 412 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).

| | | | | | |
|-----------------------|---|---------|-----------|------------------------|---|
| FT | DISULFID | 118 | 206 | BY SIMILARITY. | |
| FT | DISULFID | 125 | 197 | BY SIMILARITY. | |
| FT | DISULFID | 130 | 154 | BY SIMILARITY. | |
| FT | DISULFID | 219 | 248 | BY SIMILARITY. | |
| FT | DISULFID | 229 | 240 | BY SIMILARITY. | |
| FT | DISULFID | 297 | 330 | BY SIMILARITY. | |
| FT | DISULFID | 376 | 442 | BY SIMILARITY. | |
| FT | DISULFID | 383 | 415 | BY SIMILARITY. | |
| FT | CARBOHYD | 87 | 87 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 137 | 137 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 144 | 144 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 153 | 153 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 157 | 157 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 185 | 185 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 188 | 188 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 198 | 198 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 235 | 235 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 242 | 242 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 263 | 263 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 277 | 277 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 290 | 290 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 296 | 296 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 331 | 331 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 338 | 338 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 353 | 353 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 384 | 384 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 390 | 390 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 402 | 402 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 441 | 441 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 445 | 445 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 458 | 458 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 459 | 459 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 462 | 462 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 608 | 608 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 613 | 613 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 622 | 622 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 634 | 634 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 671 | 671 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| SQ | SEQUENCE | 853 AA; | 97043 MW; | 849B0B8CBAPF7008 | CRC64; |
| Query Match | | | | | 53.3%; Score 241; DB 1; Length 853; |
| Best Local Similarity | | | | | 85.7%; Pred. No. 1e-13; |
| Matches | | | | | 42; Conservative 4; Mismatches 3; Indels 0; Gaps 0; |
| Qy | 4 TWLEWDREINNYTSLIHSLEESQOQEKNEQELLELDKWSLWNNFNI 52 | | | | |
| Db | 624 TWLEWDREINNYTGLIYRLIEESQTQOQEKNEQELLELDKWSLWNNFNI 672 | | | | |

Search completed: February 25, 2004, 15:41:45
Job time : 18.2152 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:35:54 ; Search time 83.3861 Seconds
(without alignments)
321.625 Million cell updates/sec

Title: US-09-877-606-4
Perfect score: 452
Sequence: 1 NHTTWLEWREINNTSLIH.....SKYHIENFIARIKKLIGR 85

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTEMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 290 | 64.2 | 748 | 15 Q70606 | Q70606 human immun |
| 2 | 290 | 64.2 | 752 | 15 Q70604 | Q70604 human immun |
| 3 | 290 | 64.2 | 752 | 15 Q70605 | Q70605 human immun |
| 4 | 286 | 63.3 | 747 | 15 Q70607 | Q70607 human immun |
| 5 | 286 | 63.3 | 752 | 15 Q70608 | Q70608 human immun |
| 6 | 285 | 63.1 | 856 | 15 Q90SM7 | Q90SM7 human immun |
| 7 | 282 | 62.4 | 852 | 15 Q89797 | Q89797 human immun |
| 8 | 280 | 61.9 | 856 | 15 Q92877 | Q92877 simian-huma |
| 9 | 277 | 61.3 | 645 | 15 Q993A6 | Q993A6 human immun |
| 10 | 277 | 61.3 | 851 | 15 Q78243 | Q78243 human immun |
| 11 | 277 | 61.3 | 854 | 15 Q85582 | Q85582 human immun |
| 12 | 277 | 61.3 | 854 | 15 Q72502 | Q72502 human immun |
| 13 | 277 | 61.3 | 856 | 15 Q74090 | Q74090 human immun |
| 14 | 277 | 61.3 | 856 | 15 Q74599 | Q74599 human immun |
| 15 | 273 | 60.4 | 854 | 15 Q90178 | Q90178 human immun |
| 16 | 273 | 60.4 | 854 | 15 Q78705 | Q78705 human immun |

| | | | | | |
|----|-----|------|-----|-----------|--------------------|
| 17 | 270 | 59.7 | 757 | 15 Q9Q722 | Q9Q722 human immun |
| 18 | 265 | 58.6 | 851 | 15 O56110 | O56110 human immun |
| 19 | 265 | 58.6 | 855 | 15 Q9E1R7 | Q9E1R7 human immun |
| 20 | 264 | 58.4 | 848 | 15 O69990 | O69990 human immun |
| 21 | 263 | 58.2 | 852 | 15 Q7SVL5 | Q7SVL5 human immun |
| 22 | 262 | 58.0 | 727 | 15 Q9Q723 | Q9Q723 human immun |
| 23 | 262 | 58.0 | 847 | 15 Q7ZB20 | Q7ZB20 human immun |
| 24 | 261 | 57.7 | 358 | 15 Q78120 | Q78120 human immun |
| 25 | 261 | 57.7 | 616 | 15 Q993B0 | Q993B0 human immun |
| 26 | 261 | 57.7 | 618 | 15 Q993B2 | Q993B2 human immun |
| 27 | 261 | 57.7 | 854 | 15 Q78225 | Q78225 human immun |
| 28 | 260 | 57.5 | 838 | 15 Q03806 | Q03806 human immun |
| 29 | 260 | 57.5 | 848 | 15 Q69988 | Q69988 human immun |
| 30 | 260 | 57.5 | 855 | 15 Q03805 | Q03805 human immun |
| 31 | 260 | 57.5 | 859 | 15 Q03805 | Q03805 human immun |
| 32 | 260 | 57.5 | 862 | 15 Q9E1S2 | Q9E1S2 human immun |
| 33 | 259 | 57.3 | 635 | 15 Q90U82 | Q90U82 human immun |
| 34 | 259 | 57.3 | 850 | 15 Q7SVL4 | Q7SVL4 human immun |
| 35 | 259 | 57.3 | 852 | 15 Q69992 | Q69992 human immun |
| 36 | 259 | 57.3 | 860 | 15 Q7SVL7 | Q7SVL7 human immun |
| 37 | 257 | 56.9 | 856 | 15 Q72993 | Q72993 human immun |
| 38 | 257 | 56.9 | 859 | 15 Q80863 | Q80863 human immun |
| 39 | 257 | 56.9 | 859 | 15 Q7SUS9 | Q7SUS9 human immun |
| 40 | 257 | 56.9 | 863 | 15 Q9WJU4 | Q9WJU4 human immun |
| 41 | 257 | 56.9 | 864 | 15 Q9YP48 | Q9YP48 human immun |
| 42 | 256 | 56.6 | 851 | 15 Q8Q852 | Q8Q852 human immun |
| 43 | 256 | 56.6 | 854 | 15 O56112 | O56112 human immun |
| 44 | 256 | 56.6 | 863 | 15 Q9WJU8 | Q9WJU8 human immun |
| 45 | 256 | 56.6 | 863 | 15 O42031 | O42031 human immun |

ALIGNMENTS

RESULT 1

Q70606 PRELIMINARY; PRT; 748 AA.
ID Q70606;
AC Q70606;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "viral variability and serum antibody response in a laboratory worker
RT infected with HIV type 1 (HTLV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LW881;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
EMBL; U12032; AAA76668.1; -.
PIR; A53591; A53591.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON TER 748 748
SQ SEQUENCE 748 AA; 84224 MW; 56BBD186C67694B CRC64;


```
SQ SEQUENCE 747 AA; 84250 MW; 732E836A52245F14 CRC64;
Query Match 63.3%; Score 286; DB 15; Length 747;
Best Local Similarity 98.1%; Pred. No. 2e-16;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWNI 52
|||||
Db 619 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWNI 670

RESULT 5
Q70608 PRELIMINARY; PRT; 752 AA.
AC Q70608;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LW87-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
RT infected with HIV type 1 (HTLV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LW87-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12035; AAA76670.1; -.
DR PIR; A53591; A53591.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env.GP41.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
FT NON_TER 752 752
SQ SEQUENCE 752 AA; 84780 MW; 708672A2D0C0E8F8 CRC64;

Query Match 63.3%; Score 286; DB 15; Length 752;
Best Local Similarity 98.1%; Pred. No. 2e-16;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWNI 52
|||||
Db 624 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWNI 675

RESULT 6
Q90SM7 PRELIMINARY; PRT; 856 AA.
AC Q90SM7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=HXB2;
RA Ataman-Onal Y., Cheynet V., Verrier B.;
RT "Mutations and transcriptional alterations associated with the
RT downregulation of HIV-1 envelope glycoprotein expression following
RT acute cytopathic effects.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF358141; AAK49977.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env.GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 856 AA; 97126 MW; 9458D02B2FD734B3 CRC64;

Query Match 63.1%; Score 285; DB 15; Length 856;
Best Local Similarity 96.2%; Pred. No. 2.8e-16;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWNI 52
|||||
Db 624 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWNI 675

RESULT 7
Q89797 PRELIMINARY; PRT; 852 AA.
AC Q89797;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LW90-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
RT infected with HIV type 1 (HTLV type IIIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LW90-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12053; AAA76685.1; -.
DR EMBL; U12036; AAA76671.1; -.
DR PIR; A53591; A53591.
DR PDB; 1F23; 20-JUN-01.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env.GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 852 AA; 96347 MW; 4E865229D4EB33CF CRC64;

Query Match 62.4%; Score 282; DB 15; Length 852;
```

```

Best Local Similarity 96.2%; Pred. No. 5e-16;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
DB 620 NRTTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 671

RESULT 8
ID Q92877 PRELIMINARY; PRT; 856 AA.
AC Q92877;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Simian-Human immunodeficiency virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=57667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9098984; PubMed=9882298;
RA Cayabyab M., Karlsson G.B., Etemad-Moghadam B.A., Hofmann W.,
RA Steenbeke T., Halloran M., Fantom J.W., Axthelm M.K., Letvin N.L.,
RA Sodroski J.G.;
RT "Changes in human immunodeficiency virus type 1 envelope glycoproteins
RT responsible for the pathogenicity of a multiply passaged simian-human
RT immunodeficiency virus (SHIV-HXBc2).";
RL J. Virol. 73:976-984(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Cayabyab M., Karlsson G.B., Etemad-Moghadam B., Hofmann W.,
RA Halloran M., Axthelm M.W., Letvin N.L., Sodroski J.G.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF041850; AAD12142.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env.GP41.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 856 AA; 97151 MW; C50BE0388F873659 CRC64;

Query Match 61.9%; Score 280; DB 15; Length 856;
Best Local Similarity 63.5%; Pred. No. 7.4e-16;
Matches 54; Conservative 9; Mismatches 10; Indels 12; Gaps 1;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 60
DB 624 NHTTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 673
QY 61 IEEILSKVHIENIARIKKLIGER 85
DB 674 --DITNWLWYIKLFIMVGLVGLR 696

RESULT 9
Q993A6 PRELIMINARY; PRT; 645 AA.
AC Q993A6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Truncated envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=1007;
RX MEDLINE=21192672; PubMed=11287644;
RA Surman S., Lockett T.D., Slobod K.S., Jones B., Riberty J.M.,
RA White S.W., Doherty P.C., Hurwitz J.L.;
RT "Localization of CD4+ T cell epitope hotspots to exposed strands of
RT HIV envelope glycoprotein suggests structural influences on antigen
RT processing.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4587-4592(2001).
DR EMBL; AF321563; AAK1810.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env.GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
FT NON_TER 1
SQ SEQUENCE 645 AA; 72485 MW; B076514BE93362EC CRC64;

Query Match 61.3%; Score 277; DB 15; Length 645;
Best Local Similarity 94.2%; Pred. No. 1e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 52
DB 593 NNTTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWASLWNWFI 644

RESULT 10
Q78243 PRELIMINARY; PRT; 851 AA.
AC Q78243;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Carlini F., Federic M., Equestre M., Ricci S., Ratti G., Zibai Q.,
RA Verani P., Rossi G.B.;
RT "Sequence analysis of HIV-1 proviral DNA from a non producer
RT chronically infected HUT-78 cellular clone.";
RL J. Viral Diseases 1:40-55(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352106; PubMed=2765297;
RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,
RA Macchi B., Mangiano N., Verani P., Rossi G.;
RT "Biological and molecular characterization of producer and non
RT producer clones from HUT-78 infected with a patient HIV isolate.";
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,
RA Borsetti A., Saggio I., Verani P., Rossi G.;
RT "Variability of HIV-1 virus: Characteristics of an infected but not
RT productive clone.";
RL Int. J. Immunopharmacol. 3:17-23(1990).
DR EMBL; Z11530; CAA77628.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR PIR; S33985; S33985.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

```

```
DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 851 AA; 96630 MW; 1A3767B9B7898027 CRC64;

Query Match      61.3%; Score 277; DB 15; Length 851;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 52
Db 619 NNMTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 670

RESULT 11
ID Q85582 PRELIMINARY; PRT; 854 AA.
AC Q85582;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope polyprotein.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86281827; PubMed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
  Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
  retrovirus in human and nonhuman cells transfected with an infectious
  molecular clone.";
RT J. Virol. 59:284-291(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86281827; PubMed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
  Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
  retrovirus in human and nonhuman cells transfected with an infectious
  molecular clone.";
RT J. Virol. 59:284-291(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX Buckler C.E.;
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
  Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219406; PubMed=1373204;
RA Dai L.C., Littau R., Takahashi K., Ennis F.A.;
RT "Mutation of human immunodeficiency virus type 1 at amino acid 585 on
  gp41 results in loss of killing by CD8+ A24-restricted cytotoxic T
  lymphocytes.";
RT J. Virol. 66:3151-3154(1992).
DR EMBL; M19921; AAA44992.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97124 MW; ABA0A1BF36D5595A CRC64;

Query Match      61.3%; Score 277; DB 15; Length 854;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 52
Db 619 NNMTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 670

RESULT 12
ID Q72502 PRELIMINARY; PRT; 854 AA.
AC Q72502;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env polyprotein.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NL4-3;
RX MEDLINE=96036482; PubMed=7483282;
RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,
  Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;
RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes
  from primary virus cultures using the polymerase chain reaction.";
RT Virology 213:80-86(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NL4-3;
RX MEDLINE=86281827; PubMed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
  Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
  retrovirus in human and nonhuman cells transfected with an infectious
  molecular clone.";
RT J. Virol. 59:284-291(1986).
DR EMBL; U26942; AAB60578.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT CONFLICT 214 214 H -> L (IN REF. 2).
FT CONFLICT 530 530 A -> S (IN REF. 2).
FT CONFLICT 739 739 G -> D (IN REF. 2).
SQ SEQUENCE 854 AA; 97005 MW; PP2264B3841D1220 CRC64;

Query Match      61.3%; Score 277; DB 15; Length 854;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 52
Db 622 NNMTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 673

RESULT 13
ID Q74090 PRELIMINARY; PRT; 856 AA.
AC Q74090;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
```

```
Db 622 NNMTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 673

RESULT 12
ID Q72502 PRELIMINARY; PRT; 854 AA.
AC Q72502;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env polyprotein.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NL4-3;
RX MEDLINE=96036482; PubMed=7483282;
RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,
  Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;
RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes
  from primary virus cultures using the polymerase chain reaction.";
RT Virology 213:80-86(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NL4-3;
RX MEDLINE=86281827; PubMed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
  Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
  retrovirus in human and nonhuman cells transfected with an infectious
  molecular clone.";
RT J. Virol. 59:284-291(1986).
DR EMBL; U26942; AAB60578.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT CONFLICT 214 214 H -> L (IN REF. 2).
FT CONFLICT 530 530 A -> S (IN REF. 2).
FT CONFLICT 739 739 G -> D (IN REF. 2).
SQ SEQUENCE 854 AA; 97005 MW; PP2264B3841D1220 CRC64;

Query Match      61.3%; Score 277; DB 15; Length 854;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 52
Db 622 NNMTWMEWDREINNYTSLIHSLEESQOQKNEQELLELDKWSLWNWNI 673

RESULT 13
ID Q74090 PRELIMINARY; PRT; 856 AA.
AC Q74090;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=PM213;
RX MEDLINE=90101366; PubMed=1688473;
RA Cloyd M.W., Moore B.E.;
RT "Spectrum of Biological Properties of Human Immunodeficiency Virus
RL (HIV-1) Isolates.";
RL Virology 174:103-116(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=PM213;
RA Iwatani Y.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86068; BAA13003.1; -.
DR PIR; A53591; A53591.
DR PDB; 1F23; 20-JUN-01.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 856 AA; 97396 MW; FE3E784C423C108C CRC64;

Query Match 61.3%; Score 277; DB 15; Length 856;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 52
Db 624 NMTWMEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 675

RESULT 14
QY4599 ID Q74599 PRELIMINARY; PRT; 856 AA.
AC Q74599;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=MCKJ;
RX MEDLINE=90101366; PubMed=1688473;
RA Cloyd M.W., Moore B.E.;
RT "Spectrum of Biological Properties of Human Immunodeficiency Virus
RL (HIV-1) Isolates.";
RL Virology 174:103-116(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=MCKJ;
RA Iwatani Y.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86068; BAA12995.1; -.
DR PIR; A53591; A53591.
DR PDB; 1F23; 20-JUN-01.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 856 AA; 97396 MW; FE3E784C423C108C CRC64;

Query Match 61.3%; Score 277; DB 15; Length 856;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 52
Db 624 NMTWMEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 675

RESULT 14
QY4599 ID Q74599 PRELIMINARY; PRT; 856 AA.
AC Q74599;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Env.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=MCKJ;
RX MEDLINE=90101366; PubMed=1688473;
RA Cloyd M.W., Moore B.E.;
RT "Spectrum of Biological Properties of Human Immunodeficiency Virus
RL (HIV-1) Isolates.";
RL Virology 174:103-116(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=MCKJ;
RA Iwatani Y.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86068; BAA12995.1; -.
DR PIR; A53591; A53591.
DR PDB; 1F23; 20-JUN-01.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 856 AA; 97396 MW; FE3E784C423C108C CRC64;

Query Match 61.3%; Score 277; DB 15; Length 856;
Best Local Similarity 94.2%; Pred. No. 1.3e-15;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 52
Db 624 NMTWMEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 675

RESULT 15
QY4599 ID Q90178 PRELIMINARY; PRT; 854 AA.
AC Q90178;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95074930; PubMed=7983770;
RA Fang H., Pincus S.H.;
RT "Unique insertion sequence and pattern of CD4 expression in variants
RT selected with immunotoxins from human immunodeficiency virus type 1-
RT infected T cells.";
RL J. Virol. 69:75-81(1995).
RN [2]
RN SEQUENCE FROM N.A.
RA Fang H., Pincus S.H.;
RT "Spontaneous activation of human immunodeficiency virus type 1 in an
RT immunotoxin-resistant variant T cell line.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070521; AAC28452.1; -.
DR PIR; A53591; A53591.
DR PIR; S13288; S13288.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 96885 MW; 0808F3AED27C693B CRC64;

Query Match 60.4%; Score 273; DB 15; Length 854;
Best Local Similarity 92.3%; Pred. No. 2.9e-15;
Matches 48; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 52
Db 622 NMTWMEWDREINNTSLIHSLSIESQKQKNEQELLELDKWSLWNWNI 673

Search completed: February 25, 2004, 15:44:37
Job time : 83.3861 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:09 ; Search time 10.3228 Seconds
(without alignments)
191.599 Million cell updates/sec

Title: US-09-877-606-5

Perfect score: 48

Sequence: 1 LWNWFDI 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 48 | 100.0 | 8 | ABP15182 | Abp15182 HIV A24 s |
| 2 | 48 | 100.0 | 8 | ABP18560 | Abp18560 HIV B62 s |
| 3 | 48 | 100.0 | 8 | ABP12066 | Abp12066 HIV A02 s |
| 4 | 48 | 100.0 | 9 | ABP12213 | Abp12213 HIV A02 s |
| 5 | 48 | 100.0 | 9 | ABP17691 | Abp17691 HIV B58 s |
| 6 | 48 | 100.0 | 10 | ABP17694 | Abp17694 HIV B58 s |
| 7 | 48 | 100.0 | 10 | ABP12373 | Abp12373 HIV A02 s |
| 8 | 48 | 100.0 | 11 | ABP12500 | Abp12500 HIV A02 s |
| 9 | 48 | 100.0 | 11 | ABP139664 | Abp139664 HIV-1 gp4 |
| 10 | 48 | 100.0 | 15 | AAW25858 | AAW25858 HIV-1 T/B |
| 11 | 48 | 100.0 | 15 | AAW67374 | AAW67374 HIV-1 pep |
| 12 | 48 | 100.0 | 15 | AAW99908 | AAW99908 HIV-1 vac |
| 13 | 48 | 100.0 | 15 | AAW39707 | AAW39707 HIV1 chim |
| 14 | 48 | 100.0 | 18 | AAW68686 | AAW68686 CLTB-94 B |
| 15 | 48 | 100.0 | 18 | AAW25856 | AAW25856 HIV-1 T/B |
| 16 | 48 | 100.0 | 18 | AAW67372 | AAW67372 HIV-1 pep |
| 17 | 48 | 100.0 | 18 | AAW99907 | AAW99907 HIV-1 vac |
| 18 | 48 | 100.0 | 18 | AAW39705 | AAW39705 HIV1 chim |
| 19 | 48 | 100.0 | 20 | AAW68684 | AAW68684 CLTB-92A |
| 20 | 48 | 100.0 | 20 | AAW08066 | AAW08066 HIV pep-i |
| 21 | 48 | 100.0 | 20 | AAW25854 | AAW25854 HIV-1 T/B |
| 22 | 48 | 100.0 | 20 | AAW67370 | AAW67370 HIV-1 pep |
| 23 | 48 | 100.0 | 20 | AAW99906 | AAW99906 HIV-1 vac |
| 24 | 48 | 100.0 | 20 | AAW39703 | AAW39703 HIV1 chim |
| 25 | 48 | 100.0 | 20 | AAW39703 | AAW39703 HIV1 chim |

ALIGNMENTS

RESULT 1

ABP15182

ID ABP15182 standard; peptide; 8 AA.

XX AC ABP15182;

XX DT 11-SEP-2003 (revised)

XX DT 15-JUL-2002 (first entry)

XX DE HIV A24 super motif env peptide #62.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX PI Baker DM, Cellis E, Kubo RT, Grey RM;

XX WPI; 2001-354887/37.

XX PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1) peptide groups, useful for vaccinating against HIV-1.

XX PS Claim 32; Page 180; 448pp; English.

XX CC The present invention describes a composition (I) comprising a prepared human immunodeficiency virus-1 (HIV-1) group comprising an amino acid sequence selected from 51 defined amino acid sequences (ABP25347 to ABP25397). (I) has virucide activity and can be used in vaccines. (I) may be used for immunising subjects against HIV-1 infections. The use of group-based vaccines has several advantages over traditional vaccines, particularly when compared to the use of whole antigens in vaccine compositions. There is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. The groups for inclusion in an group-based vaccine may be selected from conserved regions of viral or

CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 48; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDDI 7
 |||||
 Db 2 LMNWFDDI 8

RESULT 2
 ABP18560
 ID ABP18560 standard; peptide; 8 AA.

XX AC ABP18560;

XX DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)

XX DE HIV B62 super motif env peptide #135.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Cheenut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.

XX PS Claim 32; Page 249; 448pp; English.

XX CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present

CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 48; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDDI 7
 |||||
 Db 2 LMNWFDDI 8

RESULT 3
 ABP12066

ID ABP12066 standard; peptide; 8 AA.

XX AC ABP12066;

XX DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)

XX DE HIV A02 super motif env peptide #123.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Cheenut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.

XX PS Claim 32; Page 116; 448pp; English.

XX CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to

CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX SQ Sequence 8 AA;

Query Match 100.0%; Score 48; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFEDI 7
 |||||
 Db 2 LNNWFEDI 8

RESULT 4
 ABP12213
 ID ABP12213 standard; peptide; 9 AA.

XX AC ABP12213;

XX DT 11-SEP-2003 (revised)

XX DT 15-JUL-2002 (first entry)

XX DE HIV A02 super motif env peptide #270.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX PI Baker DM, Cellis E, Kubo RT, Grey HM;

XX DR WPI; 2001-354887/37.

XX PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 XX peptide groups, useful for vaccinating against HIV-1.

XX PS Claim 32; Page 119; 448pp; English.

XX CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced

CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFEDI 7
 |||||
 Db 2 LNNWFEDI 8

RESULT 5

ABP17691

ID ABP17691 standard; peptide; 9 AA.

XX AC ABP17691;

XX DT 11-SEP-2003 (revised)

XX DT 15-JUL-2002 (first entry)

XX DE HIV B58 super motif env peptide #92.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX PI Baker DM, Cellis E, Kubo RT, Grey HM;

XX DR WPI; 2001-354887/37.

XX PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 XX peptide groups, useful for vaccinating against HIV-1.

XX PS Claim 32; Page 231; 448pp; English.

XX CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response

CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 48; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 |||||
 3 LWNWFDI 9

DB

RESULT 6
 ABP17694
 ID ABP17694 standard; peptide; 10 AA.
 XX
 AC ABP17694;
 XX
 DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX
 DE HIV B58 super motif env peptide #95.
 XX
 KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO200124810-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-US027766.
 XX
 PR 05-OCT-1999; 99US-00412863.
 XX
 PA (EPIM-) EPIMMUNE INC.
 XX
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX
 WPI; 2001-354887/37.
 XX
 Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.
 XX
 PS Claim 32; Page 231; 448pp; English.

CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines.
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present

CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 48; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 |||||
 4 LWNWFDI 10

DB

RESULT 7
 ABP12373
 ID ABP12373 standard; peptide; 10 AA.
 XX
 AC ABP12373;
 XX
 DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX
 DE HIV A02 super motif env peptide #430.
 XX
 KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO200124810-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-US027766.
 XX
 PR 05-OCT-1999; 99US-00412863.
 XX
 PA (EPIM-) EPIMMUNE INC.
 XX
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX
 WPI; 2001-354887/37.
 XX
 Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.
 XX
 PS Claim 32; Page 122; 448pp; English.

CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines.
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

SQ Sequence 10 AA;
 Query Match 100.0%; Score 48; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 4 LWNWFDI 10
 |||||

RESULT 8
 ABP12500
 ID ABP12500 standard; peptide; 11 AA.
 XX
 AC ABP12500;
 XX
 DT 11-SEP-2003 (revised)
 DT 15-JUL-2002 (first entry)
 XX
 DE HIV A02 super motif env peptide #557.
 XX
 KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
 KW vaccine; HIV infection; immunisation; virucide.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO200124810-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-US027766.
 XX
 PR 05-OCT-1999; 99US-00412863.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX
 PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Cellis E, Kubo RT, Grey HM;
 XX
 DR WPI; 2001-354887/37.
 XX
 PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
 PT peptide groups, useful for vaccinating against HIV-1.
 XX
 PS Claim 32; Page 125; 448pp; English.
 XX
 CC The present invention describes a composition (I) comprising a prepared
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
 CC sequence selected from 51 defined amino acid sequences (ABP25347 to
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
 CC be used for immunising subjects against HIV-1 infections. The use of
 CC group-based vaccines has several advantages over traditional vaccines,
 CC particularly when compared to the use of whole antigens in vaccine
 CC compositions. There is evidence that the immune response to whole
 CC antigens is directed largely toward variable regions of the antigen,
 CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-based vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines. An
 CC additional advantage of an group-based vaccine approach is the ability to
 CC combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABP11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

Sequence 11 AA;

Query Match 100.0%; Score 48; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 4 LWNWFDI 10
 |||||

RESULT 9
 ABR39664
 ID ABR39664 standard; peptide; 11 AA.
 XX
 AC ABR39664;
 XX
 DT 23-OCT-2003 (revised)
 DT 23-JUN-2003 (first entry)
 XX
 DE HIV-1 gp41 peptide fragment (residues 670-680).
 XX
 KW HIV-1; immune response; gp41; anti-HIV; vaccine; mAb 4E10-IgG1;
 KW anti-idiopathic antibody.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO2003022879-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 09-SEP-2002; 2002WO-BP010070.
 XX
 PR 07-SEP-2001; 2001US-0318091P.
 XX
 PA (POLY-) POLYMUN SCI IMMUNOBIOLOGISCHE FORSCHUNG.
 XX
 PI Stiegler G, Kunert R, Katinger H;
 XX
 DR WPI; 2003-354526/33.
 XX
 PT New peptide that interferes with HIV-1 entry into target cells and
 PT preferably induces an HIV-1 neutralizing immune response, where the
 PT peptide is a fragment of gp41 of HIV-1, useful for inhibiting or
 PT preventing HIV-1 infection.
 XX
 PS Claim 1; Page 21; 33pp; English.
 XX
 CC The invention relates to a peptide that interferes with HIV-1 entry into
 CC target cells and preferably induces an HIV-1 neutralizing immune
 CC response, where the peptide is a fragment of gp41 of HIV-1. The peptide,
 CC antibody, composition and vaccine are useful for inhibiting or preventing
 CC HIV-1 infection. The mAb 4E10-IgG1 is useful for eliciting or screening
 CC for an anti-idiopathic antibody that is reactive with the 4E10 binding
 CC paratope of mAb 4E10-IgG1, and that preferably mimics a fragment of gp41
 CC of HIV. The present sequence represents a peptide fragment of HIV-1 gp41
 CC of TCIA isolate HTVL IIIMN. (Updated on 23-OCT-2003 to standardise OS
 CC field)

Sequence 11 AA;

Query Match 100.0%; Score 48; DB 6; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 1 LWNWFDI 7
 |||||

RESULT 10
 AAR68688
 ID AAR68688 standard; peptide; 15 AA.
 XX
 AC AAR68688;

XX 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 07-SEP-1995 (first entry)
 XX CLTB-96 B-cell epitope.
 XX T-cell; epitope; HIV-1; core protein; p24E; B-cell; antigen; gp160; gag;
 KW pol; vaccine; multimeric peptide; AIDS; 3D organisation.
 XX Human immunodeficiency virus 1.
 XX WO9429339-A1.
 XX 22-DEC-1994.
 XX 08-JUN-1994; 94WO-CA000317.
 XX 09-JUN-1993; 93US-00073378.
 XX (CONN-) CONNAUGHT LAB LTD.
 XX Sia CDY, Chong P, Klein MH;
 XX WPI; 1995-036400/05.
 XX Novel tandem synthetic HIV-1 peptide(s) - comprising T-cell epitope of
 PT gag protein linked to B-cell epitope of V3 loop protein of an HIV-1
 PT isolate.
 XX Claim 4; Page 42; 69pp; English.
 XX This sequence represents the B-cell epitope, CLTB-96, which contains the
 CC gp41 neutralising epitope ELDKWA. This peptide is pref. linked to the T-
 CC cell epitope, p24E, in the production of a chimeric peptide. Chimeric
 CC peptides such as this are recognised by monoclonal antibody 2F5. These
 CC chimeric peptides may then be used in the production of HIV-1 vaccines.
 CC These peptide sequences may also be used in the production of multimeric
 CC peptides in which the peptides are C-terminally modified by the addition
 CC of a lys residue which is modified on its epsilon TCC amino acid to carry
 CC an additional copy of the peptide molecule. The linear and multimeric
 CC peptides may be used for the treatment of AIDS by acting to displace the
 CC binding of HIV virus to human or animal cells or by disturbing the 3D
 CC organisation of the virus. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 16-OCT-2003 to standardise OS field)
 XX SQ Sequence 15 AA;
 Query Match 100.0%; Score 48; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWNWFDI 7
 DB 8 LWNWFDI 14
 RESULT 11
 AAW25858
 ID AAW25858 standard; peptide; 15 AA.
 XX AAW25858;
 AC 25-MAR-2003 (revised)
 DT 20-OCT-1997 (first entry)
 XX HIV-1 T/B cell epitope CLTB-96 contains gp41 neutralisation epitope.
 XX HIV; human immunodeficiency virus; gag; T-cell; B-cell; epitope; env;
 KW V3 loop; vaccine; determinant; chimaeric.
 XX Synthetic.
 OS
 XX

FH Key Location/Qualifiers
 FT Region 1..6
 FT /note= "gp41 neutralisation epitope"
 XX US5639854-A.
 XX 17-JUN-1997.
 XX 09-JUN-1994; 94US-00257528.
 XX 09-JUN-1993; 93US-00073378.
 XX (CONN-) CONNAUGHT LAB LTD.
 XX Klein MH, Sia CDY, Chong P;
 XX WPI; 1997-332082/30.
 XX Tandem synthetic HIV peptide(s) useful as immunogens - comprising gag
 PT protein T-cell epitope linked to env protein B-cell epitope.
 XX Example 1; Col 21-22; 41pp; English.
 XX The invention relates to new synthetic peptides comprising at least one
 CC amino acid sequence comprising an HIV gag protein T-cell epitope linked
 CC at its C- or N-terminus to an amino acid sequence comprising a B-cell
 CC epitope of the V3 loop of an HIV env protein, which can be used to
 CC generate vaccines against HIV-1. The T-cell epitope sequence is pref.
 CC selected from the T-helper determinant core peptides p24E, p24N, p24L,
 CC p24M and p24H while the B-cell epitopes are derived from HIV strains
 CC including CLTB-56, V3MN, CLTB-29, CLTB-55, SF2, LAI, IIB, RF, 26, 2054,
 CC 1714 and BX08. The peptides are chimaeric and can be linked to a branched
 CC lys backbone. The peptides AAW25853-67 represents chimaeric T/B cell
 CC epitope peptides which also contain a gp41 neutralisation epitope.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 15 AA;
 Query Match 100.0%; Score 48; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWNWFDI 7
 DB 8 LWNWFDI 14
 RESULT 12
 AAW67374
 ID AAW67374 standard; peptide; 15 AA.
 XX AAW67374;
 AC 17-OCT-2003 (revised)
 DT 25-JAN-1999 (first entry)
 XX HIV-1 peptide epitope CLTB-96.
 DE Immunogen; vaccine; HIV-1; T-cell; B-cell; epitope; core protein; gp120;
 KW V3 loop; gp41.
 KW Human immunodeficiency virus 1.
 OS US5817754-A.
 XX 06-OCT-1998.
 XX 05-JUN-1995; 95US-00464329.
 XX 09-JUN-1993; 93US-00073378.
 PR 09-JUN-1994; 94US-00257528.
 XX (CONN-) CONNAUGHT LAB LTD.
 XX

XX Chong P, Klein MH, Sia CDY;
 XX WPI; 1998-556461/47.
 XX Synthetic human immunodeficiency virus-1 peptide(s) - containing T-cell
 PT epitope and B-cell epitope(s) are candidate vaccines against HIV-1.
 XX Disclosure; Col 21-22; 40pp; English.
 XX The invention relates to a novel immunogenic composition for use in
 CC vaccines for the treatment of HIV-1 comprising an HIV-1-derived T-cell
 CC epitope linked to an HIV-1-derived B-cell epitope. The T-cell epitopes
 CC are generally designed based on the p24 core protein and the B-cell
 CC epitopes from the V3 loop of the gp120 protein from various HIV-1
 CC strains. This peptide corresponds to an HIV-1 gp41 protein epitope. The
 CC peptide is used to generate the hybrid epitope AAW67373. (Updated on 17-
 CC OCT-2003 to standardise OS field)
 XX Sequence 15 AA;
 SQ

Query Match 100.0%; Score 48; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 8 LWNWFDI 14
 |||||

RESULT 13
 AAW99908
 ID AAW99908 standard; peptide; 15 AA.
 XX
 AC AAW99908;
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE HIV-1 vaccine synthetic peptide CLTB-96.
 XX
 KW HIV-1; human immunodeficiency virus; vaccine; T-cell epitope;
 KW gag protein; B-cell epitope; gp41 protein; chimeric; infection.
 XX
 OS Synthetic.
 OS Human immunodeficiency virus 1.
 XX
 PN US5876731-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 05-JUN-1995; 95US-00462507.
 XX
 PR 09-JUN-1993; 93US-00073378.
 PR 09-JUN-1994; 94US-00257528.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 XX
 PI Chong P, Klein MH, Sia CDY;
 XX WPI; 1999-189590/16.
 XX Synthetic chimeric HIV polypeptides - comprising gag protein T-cell
 PT epitope linked to gp41 B-cell epitope.
 XX Claim 7; Col 72; 41pp; English.
 XX

The present invention describes a synthetic peptide comprising an amino
 CC acid sequence containing a T-cell epitope of an HIV gag protein linked at
 CC its C terminus to an amino acid sequence containing a B-cell epitope of
 CC an HIV gp41 protein and containing the amino acid sequence: XILKDWX2,
 CC where X1 = E, A, G or Q, and X2 = A or T, or an amino acid sequence
 CC capable of eliciting an HIV-specific antiserum and recognizing the
 CC sequence XILKDWX2. The synthetic peptide is useful in vaccines against

CC HIV infection and in diagnostic applications. AAW98892 to AAW98906, and
 CC AAW98999 to AAW99989 represent synthetic peptides from the present
 CC invention
 XX Sequence 15 AA;
 SQ

Query Match 100.0%; Score 48; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 8 LWNWFDI 14
 |||||

RESULT 14
 AAY39707
 ID AAY39707 standard; peptide; 15 AA.
 XX
 AC AAY39707;
 XX
 DT 17-OCT-2003 (revised)
 DT 26-NOV-1999 (first entry)
 XX
 DE HIV1 chimeric peptide CLTB-96.
 XX
 KW HIV; vaccine; immunogenic composition; T cell epitope; B cell epitope;
 KW infection; antibody; antiviral.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5951986-A.
 XX
 PD 14-SEP-1999.
 XX
 PF 06-JUN-1995; 95US-00467881.
 XX
 PR 09-JUN-1993; 93US-00073378.
 PR 09-JUN-1994; 94US-00257528.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 XX
 PI Klein MH, Chong P, Sia CDY;
 XX WPI; 1999-550482/46.
 XX
 PT Immunogenic composition containing synthetic fusion polypeptides
 PT containing both the T and B cell epitopes of the human immunodeficiency
 PT virus, useful antigens in producing vaccines.
 XX
 PS Example 1; Col 23-24; 43pp; English.
 XX

This sequence represents a fragment of a HIV1 protein, and can be used in
 CC the immunogenic composition of the invention. The composition comprises a
 CC synthetic fusion polypeptide which includes a sequence encoding 1 or more
 CC T cell epitopes and a sequence encoding 1 or more B cell epitopes and a
 CC carrier. Both the T cell and B cell epitopes are derived from HIV
 CC proteins. The compositions are useful as vaccines against HIV infection.
 CC The composition induces HIV-1-specific polyclonal antibodies that are
 CC opsonising and antiviral. The peptide components may be selected to
 CC induce a response against different viral isolates and in subjects who
 CC recognise different T cell epitopes. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 48; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 8 LWNWFDI 14
 |||||

```

RESULT 15
AAR68686
ID AAR68686 standard; peptide; 18 AA.
XX
AC AAR68686;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 07-SEP-1995 (first entry)
XX
DE CLTB-94 B-cell epitope.
XX
KW T-cell; epitope; HIV-1; core protein; p24E; B-cell; antigen; gp160; gag;
KW pol; vaccine; multimeric peptide; AIDS; 3D organisation.
XX
OS Human immunodeficiency virus 1.
XX
PN WO9429339-A1.
XX
PD 22-DEC-1994.
XX
PF 08-JUN-1994; 94WO-CA000317.
XX
PR 09-JUN-1993; 93US-00073378.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Sia CDY, Chong P, Klein MH;
XX
DR WPI; 1995-036400/05.
XX
PT Novel tandem synthetic HIV-1 peptide(s) - comprising T-cell epitope of
PT gag protein linked to B-cell epitope of V3 loop protein of an HIV-1
PT isolate.
XX
PS Claim 4; Page 42; 69pp; English.
XX
CC This sequence represents the B-cell epitope, CLTB-94, which contains the
CC gp41 neutralising epitope ELDKWA. This peptide is pref. linked to the T-
CC cell epitope, p24E, in the production of a chimeric peptide. Chimeric
CC peptides such as this are recognised by monoclonal antibody 2F5. These
CC chimeric peptides may then be used in the production of HIV-1 vaccines.
CC These peptide sequences may also be used in the production of multimeric
CC peptides in which the peptides are C-terminally modified by the addition
CC of a Lys residue which is modified on its epsilon TCC amino acid to carry
CC an additional copy of the peptide molecule. The linear and multimeric
CC peptides may be used for the treatment of AIDS by acting to displace the
CC binding of HIV virus to human or animal cells or by disturbing the 3D
CC organisation of the virus. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 18 AA;

```

```

Query Match 100.0%; Score 48; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LWNWFEDI 7
   |||||
DB 11 LWNWFEDI 17

```

Search completed: February 25, 2004, 15:41:06
Job time : 12.3228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:37:04 ; Search time 2.79114 Seconds
(without alignments)
129.475 Million cell updates/sec

Title: US-09-877-606-5
Perfect score: 48
Sequence: 1 LWNWFDI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents:AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 48 | 100.0 | 15 | 1 | US-08-257-528B-59 |
| 2 | 48 | 100.0 | 15 | 1 | US-08-460-602A-59 |
| 3 | 48 | 100.0 | 15 | 1 | US-08-463-966A-59 |
| 4 | 48 | 100.0 | 15 | 1 | US-08-465-217A-59 |
| 5 | 48 | 100.0 | 15 | 2 | US-08-464-329A-59 |
| 6 | 48 | 100.0 | 15 | 2 | US-08-462-507A-59 |
| 7 | 48 | 100.0 | 15 | 2 | US-08-467-881A-59 |
| 8 | 48 | 100.0 | 18 | 1 | US-08-257-528B-57 |
| 9 | 48 | 100.0 | 18 | 1 | US-08-460-602A-57 |
| 10 | 48 | 100.0 | 18 | 1 | US-08-463-966A-57 |
| 11 | 48 | 100.0 | 18 | 1 | US-08-465-217A-57 |
| 12 | 48 | 100.0 | 18 | 2 | US-08-464-329A-57 |
| 13 | 48 | 100.0 | 18 | 2 | US-08-462-507A-57 |
| 14 | 48 | 100.0 | 18 | 2 | US-08-467-881A-57 |
| 15 | 48 | 100.0 | 20 | 1 | US-08-218-025A-86 |
| 16 | 48 | 100.0 | 20 | 1 | US-08-257-528B-55 |
| 17 | 48 | 100.0 | 20 | 1 | US-08-460-602A-55 |
| 18 | 48 | 100.0 | 20 | 1 | US-08-463-966A-55 |
| 19 | 48 | 100.0 | 20 | 1 | US-08-465-217A-55 |
| 20 | 48 | 100.0 | 20 | 2 | US-08-464-329A-55 |
| 21 | 48 | 100.0 | 20 | 2 | US-08-462-507A-55 |
| 22 | 48 | 100.0 | 20 | 2 | US-08-467-881A-55 |
| 23 | 48 | 100.0 | 30 | 1 | US-08-257-528B-58 |
| 24 | 48 | 100.0 | 30 | 1 | US-08-460-602A-58 |
| 25 | 48 | 100.0 | 30 | 1 | US-08-463-966A-58 |
| 26 | 48 | 100.0 | 30 | 1 | US-08-465-217A-58 |
| 27 | 48 | 100.0 | 30 | 2 | US-08-464-329A-58 |

| | | | | | | |
|----|----|-------|-----|---|-------------------|-------------------|
| 28 | 48 | 100.0 | 30 | 2 | US-08-462-507A-58 | Sequence 58, Appl |
| 29 | 48 | 100.0 | 30 | 2 | US-08-467-881A-58 | Sequence 58, Appl |
| 30 | 48 | 100.0 | 33 | 1 | US-08-257-528B-56 | Sequence 56, Appl |
| 31 | 48 | 100.0 | 33 | 1 | US-08-460-602A-56 | Sequence 56, Appl |
| 32 | 48 | 100.0 | 33 | 1 | US-08-463-966A-56 | Sequence 56, Appl |
| 33 | 48 | 100.0 | 33 | 1 | US-08-465-217A-56 | Sequence 56, Appl |
| 34 | 48 | 100.0 | 33 | 2 | US-08-464-329A-56 | Sequence 56, Appl |
| 35 | 48 | 100.0 | 33 | 2 | US-08-462-507A-56 | Sequence 56, Appl |
| 36 | 48 | 100.0 | 33 | 2 | US-08-467-881A-56 | Sequence 56, Appl |
| 37 | 48 | 100.0 | 35 | 1 | US-08-257-528B-54 | Sequence 54, Appl |
| 38 | 48 | 100.0 | 35 | 1 | US-08-460-602A-54 | Sequence 54, Appl |
| 39 | 48 | 100.0 | 35 | 1 | US-08-463-966A-54 | Sequence 54, Appl |
| 40 | 48 | 100.0 | 35 | 1 | US-08-465-217A-54 | Sequence 54, Appl |
| 41 | 48 | 100.0 | 35 | 2 | US-08-464-329A-54 | Sequence 54, Appl |
| 42 | 48 | 100.0 | 35 | 2 | US-08-462-507A-54 | Sequence 54, Appl |
| 43 | 48 | 100.0 | 35 | 2 | US-08-467-881A-54 | Sequence 54, Appl |
| 44 | 48 | 100.0 | 138 | 4 | US-09-570-921-8 | Sequence 8, Appl |
| 45 | 48 | 100.0 | 138 | 4 | US-09-570-921-14 | Sequence 14, Appl |

ALIGNMENTS

RESULT 1
US-08-257-528B-59
; Sequence 59, Application US/08257528B
; Patent No. 5639854
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,528B
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-336 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-257-528B-59

Query Match 100.0%; Score 48; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||
Db 8 LWNWFDI 14

```
RESULT 2
US-08-460-602A-59
; Sequence 59, Application US/08460602A
; Patent No. 5759769
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,602A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-450 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-460-602A-59
Query Match 100.0%; Score 48; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 8 LWNWFDI 14

RESULT 3
US-08-463-966A-59
; Sequence 59, Application US/08463966A
; Patent No. 5795955
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
```

```
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,966A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-487 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-463-966A-59
Query Match 100.0%; Score 48; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 8 LWNWFDI 14

RESULT 4
US-08-465-217A-59
; Sequence 59, Application US/08465217A
; Patent No. 5800822
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,217A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
```

;; PRIOR APPLICATION DATA: 08/073,378
;; FILING DATE: 09-JUN-1993
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, MICHAEL I.
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-486 MIS:jb
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; INFORMATION FOR SEQ ID NO: 59:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-465-217A-59

Query Match 100.0%; Score 48; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
Db 8 LNNWFDI 14

RESULT 5
US-08-464-329A-59
; Sequence 59, Application US/08464329A
; Patent No. 5817754
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,329A
; FILING DATE: 05-JUN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/073,378
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-449 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-464-329A-59

Query Match 100.0%; Score 48; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
Db 8 LNNWFDI 14

RESULT 6
US-08-462-507A-59
; Sequence 59, Application US/08462507A
; Patent No. 5876731
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,507A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/073,378
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-451 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-462-507A-59

Query Match 100.0%; Score 48; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
Db 8 LNNWFDI 14

RESULT 7
US-08-467-881A-59

```
; Sequence 59, Application US/08467881A
; Patent No. 5951986
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,881A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; QUERY MATCH
; BEST LOCAL SIMILARITY 100.0%; Score 48; DB 2; Length 15;
; MATCHES 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 8 LWNWFDI 14

RESULT 8
US-08-257-528B-57
; Sequence 57, Application US/08257528B
; Patent No. 5639854
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
```

```
; Sequence 59, Application US/08467881A
; Patent No. 5951986
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,528B
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-336 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-257-528B-57
; QUERY MATCH
; BEST LOCAL SIMILARITY 100.0%; Score 48; DB 1; Length 18;
; MATCHES 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 11 LWNWFDI 17

RESULT 9
US-08-460-602A-57
; Sequence 57, Application US/08460602A
; Patent No. 5759769
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,602A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-450 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
```

TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-602A-57

Query Match 100.0%; Score 48; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
|||||
DB 11 LMNWFDI 17

RESULT 10
US-08-463-966A-57
Sequence 57, Application US/08463966A
Patent No. 5795955
GENERAL INFORMATION:

APPLICANT: SIA, Charles D.Y.
APPLICANT: CHONG, Pele
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,966A
FILING DATE: 05-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,528
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/073,378
FILING DATE: 09-JUN-1993

ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-487 MIS:jb
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-463-966A-57

Query Match 100.0%; Score 48; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
|||||

Db 11 LMNWFDI 17

RESULT 11
US-08-465-217A-57
Sequence 57, Application US/08465217A
Patent No. 5800822
GENERAL INFORMATION:

APPLICANT: SIA, Charles D.Y.
APPLICANT: CHONG, Pele
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,217A
FILING DATE: 05-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,528
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/073,378
FILING DATE: 09-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.

REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-486 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-465-217A-57

Query Match 100.0%; Score 48; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
|||||
DB 11 LMNWFDI 17

RESULT 12
US-08-464-329A-57
Sequence 57, Application US/08464329A
Patent No. 5817754
GENERAL INFORMATION:

APPLICANT: SIA, Charles D.Y.
APPLICANT: CHONG, Pele
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney

```

; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,507A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-449 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-464-329A-57

Query Match 100.0%; Score 48; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 11 LWNWFDI 17

RESULT 13
US-08-462-507A-57
; Sequence 57, Application US/08462507A
; Patent No. 5876731
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,507A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-488 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 57:

```

```

; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/073,378
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-451 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-462-507A-57

Query Match 100.0%; Score 48; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 11 LWNWFDI 17

RESULT 14
US-08-467-881A-57
; Sequence 57, Application US/08467881A
; Patent No. 5951986
; GENERAL INFORMATION:
; APPLICANT: SIA, Charles D.Y.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Tandem Synthetic HIV-1 Peptides
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,881A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,528
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-488 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 57:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-881A-57

Query Match 100.0%; Score 48; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNNWFDI 7
|
|
|
|
|
|
|
Db 11 LNNWFDI 17

RESULT 15

US-08-218-025A-86
; Sequence 86, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Ugen, Kenneth E.
; APPLICANT: Williams, William V.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; TITLE OF INVENTION: and Treating Certain HIV Infected Patients
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 5556744ristown Road
; City: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,025A
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/891,451
; FILING DATE: 29-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST33A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-218-025A-86

Query Match 100.0%; Score 48; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNNWFDI 7
|
|
|
|
|
|
|
Db 10 LNNWFDI 16

Search completed: February 25, 2004, 15:46:55
Job time : 3.79114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:44:50 ; Search time 5.8481 Seconds
(without alignments)
252.744 Million cell updates/sec

Title: US-09-877-606-5

Perfect score: 48

Sequence: 1 LNWVFDI 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

| | |
|-----|--|
| 1: | /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.* |
| 2: | /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.* |
| 3: | /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.* |
| 4: | /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.* |
| 5: | /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.* |
| 6: | /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.* |
| 7: | /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.* |
| 8: | /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.* |
| 9: | /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.* |
| 10: | /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.* |
| 11: | /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.* |
| 12: | /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.* |
| 13: | /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.* |
| 14: | /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.* |
| 15: | /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.* |
| 16: | /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.* |
| 17: | /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.* |
| 18: | /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 48 | 100.0 | 146 | 13 | US-10-000-321-10 |
| 2 | 48 | 100.0 | 267 | 9 | US-09-854-816-15 |
| 3 | 48 | 100.0 | 268 | 9 | US-09-854-816-8 |
| 4 | 48 | 100.0 | 268 | 9 | US-09-854-816-13 |
| 5 | 48 | 100.0 | 268 | 9 | US-09-854-816-14 |
| 6 | 48 | 100.0 | 268 | 9 | US-09-854-816-26 |
| 7 | 48 | 100.0 | 268 | 9 | US-09-854-816-35 |
| 8 | 48 | 100.0 | 268 | 9 | US-09-854-816-39 |
| 9 | 48 | 100.0 | 268 | 9 | US-09-854-816-40 |
| 10 | 48 | 100.0 | 268 | 9 | US-09-854-816-68 |
| 11 | 48 | 100.0 | 268 | 9 | US-09-854-816-70 |
| 12 | 48 | 100.0 | 268 | 9 | US-09-854-816-71 |
| 13 | 48 | 100.0 | 268 | 9 | US-09-854-816-72 |
| 14 | 48 | 100.0 | 269 | 9 | US-09-854-816-20 |
| 15 | 48 | 100.0 | 269 | 9 | US-09-854-816-21 |

| | | | | | | |
|----|----|-------|-----|----|-------------------|--------------------|
| 16 | 48 | 100.0 | 269 | 9 | US-09-854-816-22 | Sequence 22, Appl |
| 17 | 48 | 100.0 | 269 | 9 | US-09-854-816-33 | Sequence 33, Appl |
| 18 | 48 | 100.0 | 269 | 9 | US-09-854-816-34 | Sequence 34, Appl |
| 19 | 48 | 100.0 | 269 | 9 | US-09-854-816-37 | Sequence 37, Appl |
| 20 | 48 | 100.0 | 269 | 9 | US-09-854-816-96 | Sequence 96, Appl |
| 21 | 48 | 100.0 | 270 | 9 | US-09-854-816-67 | Sequence 67, Appl |
| 22 | 48 | 100.0 | 351 | 9 | US-09-886-156-46 | Sequence 46, Appl |
| 23 | 48 | 100.0 | 351 | 9 | US-09-886-150-46 | Sequence 46, Appl |
| 24 | 48 | 100.0 | 351 | 10 | US-09-886-149-46 | Sequence 46, Appl |
| 25 | 48 | 100.0 | 351 | 10 | US-09-886-159-46 | Sequence 46, Appl |
| 26 | 48 | 100.0 | 351 | 14 | US-10-326-090-46 | Sequence 46, Appl |
| 27 | 48 | 100.0 | 360 | 14 | US-10-214-670-59 | Sequence 59, Appl |
| 28 | 48 | 100.0 | 360 | 14 | US-10-214-670-60 | Sequence 60, Appl |
| 29 | 48 | 100.0 | 360 | 14 | US-10-214-670-61 | Sequence 61, Appl |
| 30 | 48 | 100.0 | 579 | 14 | US-10-032-162-15 | Sequence 15, Appl |
| 31 | 48 | 100.0 | 619 | 9 | US-09-891-609-4 | Sequence 4, Appl |
| 32 | 48 | 100.0 | 625 | 14 | US-10-032-162-17 | Sequence 17, Appl |
| 33 | 48 | 100.0 | 643 | 14 | US-10-032-162-13 | Sequence 13, Appl |
| 34 | 48 | 100.0 | 646 | 9 | US-09-891-609-2 | Sequence 2, Appl |
| 35 | 48 | 100.0 | 842 | 14 | US-10-190-435-2 | Sequence 2, Appl |
| 36 | 48 | 100.0 | 842 | 14 | US-10-241-009-2 | Sequence 2, Appl |
| 37 | 48 | 100.0 | 842 | 14 | US-10-190-434B-2 | Sequence 2, Appl |
| 38 | 48 | 100.0 | 842 | 14 | US-10-190-305A-2 | Sequence 2, Appl |
| 39 | 48 | 100.0 | 847 | 9 | US-09-476-242-2 | Sequence 2, Appl |
| 40 | 48 | 100.0 | 853 | 14 | US-10-190-435-133 | Sequence 133, Appl |
| 41 | 48 | 100.0 | 854 | 14 | US-10-369-294-16 | Sequence 16, Appl |
| 42 | 48 | 100.0 | 854 | 14 | US-10-369-294-17 | Sequence 17, Appl |
| 43 | 48 | 100.0 | 857 | 14 | US-10-190-435-132 | Sequence 132, Appl |
| 44 | 48 | 100.0 | 859 | 14 | US-10-190-435-137 | Sequence 137, Appl |
| 45 | 48 | 100.0 | 860 | 14 | US-10-190-435-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1
US-10-000-321-10
; Sequence 10, Application US/10000321
; Publication No. US20020123039A1
; GENERAL INFORMATION:
; APPLICANT: BRUST, Stefan
; KNAPP, Stefan
; GERKEN, Manfred
; GUERTLER, Lutz
; TITLE OF INVENTION: Peptides derived from a retrovirus of
; the HIV group, and their use
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/000,321
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/131,551
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 58315/106/BEAK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399

TELEX: 904136
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-000-321-10

Query Match 100.0%; Score 48; DB 13; Length 146;
Best Local Similarity 100.0%; Pred. No. 7.5; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7

Db 138 LNNWFDI 144

RESULT 2

US-09-854-816-15
; Sequence 15, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovaanik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Ph.D., Timothy E.
REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: F1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids
TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-854-816-15

Query Match 100.0%; Score 48; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7

Db 138 LNNWFDI 144

Db 198 LNNWFDI 204

RESULT 3

US-09-854-816-8
; Sequence 8, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovaanik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Ph.D., Timothy E.
REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: F1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids
TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-854-816-8

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7

Db 199 LNNWFDI 205

RESULT 4

US-09-854-816-13
; Sequence 13, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovaanik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

```

;
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-854-816-13

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
DB 199 LWNWFDI 205

RESULT 5
US-09-854-816-14
; Sequence 14, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovasnik
; James A. Wells
;
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; Making Same
;
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001

```

```

;
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-854-816-14

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
DB 199 LWNWFDI 205

RESULT 6
US-09-854-816-26
; Sequence 26, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovasnik
; James A. Wells
;
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; Making Same
;
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

```

SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-854-816-26

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 199 LWNWFDI 205

RESULT 7

US-09-854-816-35
Sequence 35, Application US/09854816
Patent No. US20020151473A1

GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovaenik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816

FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Ph.D., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 35:

US-09-854-816-35

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 199 LWNWFDI 205

RESULT 8

US-09-854-816-39

Sequence 39, Application US/09854816

Patent No. US20020151473A1

GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovaenik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816

FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Ph.D., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 39:

US-09-854-816-39

Query Match 100.0%; Score 48; DB 9; Length 268;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7

Db 199 LWNWFDI 205

RESULT 9

US-09-854-816-40

Sequence 40, Application US/09854816

Patent No. US20020151473A1

GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovaenik
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-854-816-40

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNNWFEDI 7
Db 199 LNNWFEDI 205

RESULT 10
US-09-854-816-68
Sequence 68, Application US/09854816
Patent No. US20020151473A1
GENERAL INFORMATION:
APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovasnik
James A. Wells
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-854-816-68

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNNWFEDI 7
Db 199 LNNWFEDI 205

RESULT 11
US-09-854-816-70
Sequence 70, Application US/09854816
Patent No. US20020151473A1
GENERAL INFORMATION:
APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovasnik
James A. Wells
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-854-816-70

Query Match 100.0%; Score 48; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PHD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-854-816-20

Query Match      100.0%; Score 48; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNNWFPI 7
      |||||
Db      200 LNNWFPI 206

TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-854-816-21

Query Match      100.0%; Score 48; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNNWFPI 7
      |||||
Db      200 LNNWFPI 206

Search completed: February 25, 2004, 16:04:25
Job time : 5.8481 secs
```

```
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PHD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-854-816-20

Query Match      100.0%; Score 48; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LNNWFPI 7
      |||||
Db      200 LNNWFPI 206

RESULT 15
US-09-854-816-21
; Sequence 21, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovasnik
; James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; Making Same
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PHD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:36:14 ; Search time 2.25949 Seconds
(without alignments)
298.005 Million cell updates/sec

Title: US-09-877-606-5
Perfect score: 48
Sequence: 1 LWNWFDI 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 48 | 100.0 | 357 | 2 S22006 | envelope protein g |
| 2 | 48 | 100.0 | 357 | 2 S22004 | envelope protein g |
| 3 | 48 | 100.0 | 358 | 2 S21998 | envelope protein g |
| 4 | 48 | 100.0 | 843 | 1 H44001 | env polypeptide pr |
| 5 | 48 | 100.0 | 847 | 2 T09448 | envelope glycoprot |
| 6 | 48 | 100.0 | 847 | 2 S13289 | env protein - huma |
| 7 | 48 | 100.0 | 854 | 1 VCLJJI | env polypeptide pr |
| 8 | 48 | 100.0 | 856 | 1 A44963 | env polypeptide pr |
| 9 | 48 | 100.0 | 859 | 1 VCLJMN | env polypeptide pr |
| 10 | 43 | 89.6 | 851 | 2 S33985 | env polypeptide - |
| 11 | 43 | 89.6 | 852 | 1 VCLJBR | env polypeptide - |
| 12 | 43 | 89.6 | 852 | 2 T12016 | envelope glycoprot |
| 13 | 43 | 89.6 | 853 | 2 S54384 | envelope polypept |
| 14 | 43 | 89.6 | 854 | 2 S13288 | env protein - huma |
| 15 | 43 | 89.6 | 855 | 1 VCLJZR | env polypeptide pr |
| 16 | 43 | 89.6 | 856 | 1 VCLJH3 | env polypeptide pr |
| 17 | 43 | 89.6 | 861 | 1 VCLJLV | env polypeptide pr |
| 18 | 43 | 89.6 | 861 | 1 VCLJSC | env polypeptide pr |
| 19 | 42 | 87.5 | 136 | 2 JU0266 | envelope polypept |
| 20 | 42 | 87.5 | 136 | 2 JT0954 | envelope polypept |
| 21 | 42 | 87.5 | 357 | 2 S21996 | envelope protein g |
| 22 | 42 | 87.5 | 357 | 2 S21992 | envelope protein g |
| 23 | 42 | 87.5 | 358 | 2 S22002 | envelope protein g |
| 24 | 42 | 87.5 | 358 | 2 S22000 | envelope protein g |
| 25 | 42 | 87.5 | 358 | 2 S70417 | envelope protein g |
| 26 | 42 | 87.5 | 436 | 2 G97186 | diverged glycosyl |
| 27 | 42 | 87.5 | 443 | 2 C41621 | env polypeptide p |
| 28 | 42 | 87.5 | 454 | 2 B41621 | env polypeptide D |
| 29 | 42 | 87.5 | 729 | 1 VCLJXK | env polypeptide pr |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 42 | 87.5 | 846 | 1 VCLJND | env polypeptide pr |
| 31 | 42 | 87.5 | 855 | 1 VCLJAJ | env polypeptide pr |
| 32 | 42 | 87.5 | 856 | 1 VCLJ3W | env polypeptide pr |
| 33 | 42 | 87.5 | 859 | 2 T01672 | envelope polypept |
| 34 | 42 | 87.5 | 861 | 1 VCLJKB | env polypeptide pr |
| 35 | 42 | 87.5 | 1044 | 2 H97186 | glycosyltransferas |
| 36 | 41 | 85.4 | 357 | 2 S21994 | envelope protein g |
| 37 | 41 | 85.4 | 1034 | 2 G90591 | hypothetical prote |
| 38 | 40 | 83.3 | 274 | 2 E83837 | hypothetical prote |
| 39 | 40 | 83.3 | 606 | 2 JCS604 | ABC-transporting p |
| 40 | 40 | 83.3 | 735 | 2 D86465 | probable integral |
| 41 | 40 | 83.3 | 863 | 2 AS3034 | gag polypeptide - |
| 42 | 40 | 83.3 | 877 | 2 S49197 | envelope protein p |
| 43 | 39 | 81.2 | 1114 | 2 T30299 | dynein heavy chain |
| 44 | 39 | 81.2 | 1415 | 2 C83070 | conserved hypochet |
| 45 | 38 | 79.2 | 241 | 2 S59377 | probable membrane |

ALIGNMENTS

RESULT 1

S22006

envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 4L)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S70420; S22006

R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr

A;Reference number: S70417; MUID:92144209; PMID:1736940

A;Accession: S70420

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-357 <ST2>

A;Cross-references: EMBL:X61354; NID:g60190; PIDN:CAA43620.1; PID:g60191

A;Experimental source: patient L

A;Note: submitted to the EMBL Data Library, July 1991

C;Superfamily: type E retrovirus env polypeptide

Query Match 100.0%; Score 48; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||||
Db 170 LWNWFDI 176

RESULT 2

S22004

envelope protein gp120/gp41 - human immunodeficiency virus type 1

C;Species: human immunodeficiency virus type 1, HIV-1

A;Variety: isolate 4B

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000

C;Accession: S22004; S70419

R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.

submitted to the EMBL Data Library, July 1991

A;Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determ

A;Reference number: S21990

A;Accession: S22004

A;Molecule type: DNA

A;Residues: 1-357 <STE1>

A;Cross-references: EMBL:X61353; NID:g60188; PIDN:CAA43618.1; PID:g60189

R;Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A;Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr

A;Reference number: S70417; MUID:92144209; PMID:1736940

A;Accession: S70419

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-292,'X',294-357 <STE2>

A;Cross-references: EMBL:X61353; NID:g60188

C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 48; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||

Db 170 LWNWFDI 176

RESULT 3

S21998

envelope protein gp120/gp41 - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

A:Variety: isolate 28

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000

C:Accession: S21998; S70425

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.

submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi

A:Reference number: S21990

A:Accession: S21998

A:Molecule type: DNA

A:Residues: 1-358 <STE1>

A:Cross-references: EMBL:X61359; NID:G60182; PIDN:CAA43630.1; PID:G60183

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cereb

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70425

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-222,'X',224-358 <STE2>

A:Cross-references: EMBL:X61359; NID:G60182; PIDN:CAA43630.1; PID:G60183

C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 48; DB 2; Length 358;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||

Db 171 LWNWFDI 177

RESULT 4

H44001

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)

N:Alternate names: coat polyprotein

N:Contains: coat protein gp120; coat protein gp41

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994

C:Accession: H44001

R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.

J. Virol. 66, 6587-6600, 1992

A:Title: Complete nucleotide sequence, genome organization, and biological properties of

A:Reference number: A44001; MUID:93021387; PMID:1404605

A:Accession: H44001

A:Molecule type: DNA

A:Residues: 1-843 <LIY>

A:Cross-references: GB:M93258

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F;1-29/Domain: signal sequence #status predicted <SIG>

F;19-35/Region: hydrophobic

F;30-489/Product: coat protein gp120 #status predicted <GP1>

F;490-843/Product: coat protein gp41 #status predicted <GP2>

F;499-515/Region: hydrophobic

F;673-689/Region: hydrophobic

F:738-755/Domain: transmembrane #status predicted <TMN>
F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,431;

Query Match 100.0%; Score 48; DB 1; Length 843;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||

Db 656 LWNWFDI 662

RESULT 5

T09448

envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999

C:Accession: T09448

R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,

submitted to the EMBL Data Library, July 1996

A:Reference number: Z16673

A:Accession: T09448

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-847 <PAN>

A:Cross-references: EMBL:U63632; NID:gl465777; PID:gl465781

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 48; DB 2; Length 847;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||

Db 660 LWNWFDI 666

RESULT 6

S13289

env protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997

C:Accession: S13289

R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.

Nature 348, 69-73, 1990

A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120

A:Reference number: S13288; MUID:91043044; PMID:2172833

A:Accession: S13289

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-847 <OBR>

C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 48; DB 2; Length 847;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
|||||

Db 660 LWNWFDI 666

RESULT 7

VCLJ51

env polyprotein precursor - simian immunodeficiency virus SIVcpz

N:Alternate names: coat polyprotein

N:Contains: coat protein gp120; coat protein gp41

C:Species: simian immunodeficiency virus SIVcpz

A:Note: host Pan troglodytes (chimpanzee)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C:Accession: S09990

R;Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A;Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A;Reference number: S09983; MUID:90259077; PMID:2188136
A;Accession: S09990
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-854 <HUE>
A;Cross-references: EMBL:X52154; NID:G58866; PIDN:CAA36407.1; PID:G58874
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;1-500/Product: coat protein gp120 #status predicted <CP1>
F;1-854/Product: coat protein gp41 #status predicted <CP2>
F;501-517/Domain: transmembrane #status predicted <TM1>
F;517-693/Domain: transmembrane #status predicted <TM2>
F;693-821/Domain: transmembrane #status predicted <TM3>
F;821-143,154,158,186,195,239,260,267,274,299,331,336,351,356,384,392,426,432,446,449
Query Match 100.0%; Score 48; DB 1; Length 854;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMNWFDI 7
Db 660 LMNWFDI 666
RESULT 8
A44963
env polyprotein precursor - human immunodeficiency virus type 1 (isolate Z321)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 28-May-1999
C;Accession: A44963
R;Srinivasan, A.; York, D.; Butler Jr., D.; Jannoun-Nasr, R.; Getchell, J.; McCormick, J.
AIDS Res. Hum. Retroviruses 5, 121-129, 1989
A;Title: Molecular characterization of HIV-1 isolated from a serum collected in 1976: nu
A;Reference number: A44963; MUID:89228766; PMID:2713163
A;Accession: A44963
A;Molecule type: DNA
A;Residues: 1-856 <SRI>
A;Cross-references: GB:M15896; NID:G329392; PIDN:AAB53948.1; PID:G329394
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
F;1-29/Domain: signal sequence #status predicted <SIG>
F;1-520/Product: coat protein gp120 #status predicted <CP1>
F;521-856/Product: coat protein gp41 #status predicted <CP2>
F;854-705/Domain: transmembrane #status predicted <TMN>
F;87,132,138,152,156,183,198,242,263,277,294,302,339,393,398,402,411,448,461,462,465,611
Query Match 100.0%; Score 48; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMNWFDI 7
Db 669 LMNWFDI 675
RESULT 9
VCLJMN
env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997

C;Accession: A28922
R;Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-St
Virology 164, 531-536, 1988
A;Title: Envelope sequences of two new United States HIV-1 isolates.
A;Reference number: A28922; MUID:88219542; PMID:3369091
A;Accession: A28922
A;Molecule type: DNA
A;Residues: 1-859 <GUR>
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prot
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-859/Product: env polyprotein #status predicted <EP1>
F;87,129,135,140,141,146,161,165,191,202,246,267,281,294,300,336,343,359,365,389,395,40
Query Match 100.0%; Score 48; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMNWFDI 7
Db 673 LMNWFDI 679
RESULT 10
S33985
env polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C;Accession: S33985
R;Carlini, F.
submitted to the EMBL Data Library, November 1991
A;Reference number: S33979
A;Accession: S33985
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-851 <CAR>
A;Cross-references: EMBL:211530; NID:G60192; PIDN:CAA77628.1; PID:G60199
C;Superfamily: type E retrovirus env polyprotein
Query Match 89.6%; Score 43; DB 2; Length 851;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMNWFDI 7
Db 664 LMNWFDI 670
RESULT 11
VCLJBR
env polyprotein - human immunodeficiency virus type 1 (isolate BR)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997
C;Accession: A31667
R;Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar,
Virology 168, 79-89, 1989
A;Title: Biological and molecular characterization of human immunodeficiency virus (HIV
A;Reference number: A94389; MUID:89085613; PMID:2789516
A;Accession: A31667
A;Molecule type: DNA
A;Residues: 1-852 <ANA>
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; polyprotein; transmembrane protein
F;1-516/Product: coat protein gp120 #status predicted <CP1>
F;517-852/Product: coat protein gp41 #status predicted <CP2>
Query Match 89.6%; Score 43; DB 1; Length 852;
Best Local Similarity 85.7%; Pred. No. 22;

```
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 665 LWNWFNI 671
|||||:|

RESULT 12
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain scl4.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g2351784
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.6%; Score 43; DB 2; Length 852;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 665 LWSWFDI 671
|||||:|

RESULT 13
S54384
envelope polyprotein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
A:Accession: S54384
R:Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A:Reference number: S54377
A:Accession: S54384
A>Status: preliminary
A:Molecule type: Genomic RNA
A:Residues: 1-853 <THE>
A:Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: polyprotein

Query Match 89.6%; Score 43; DB 2; Length 853;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 666 LWNWFNI 672
|||||:|

RESULT 14
S13288
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13288
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13288
A>Status: preliminary
```

```
A:Molecule type: DNA
A:Residues: 1-854 <ORF>
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.6%; Score 43; DB 2; Length 854;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 667 LWNWFNI 673
|||||:|

RESULT 15
VCLJZR
env polyprotein precursor - human immunodeficiency virus 2r-6
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus 2r-6
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C:Accession: D26192
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; C
Gene 52, 71-82, 1987
A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot
A:Reference number: A26192; MUID:87248097; PMID:3036660
A:Accession: D26192
A:Molecule type: DNA
A:Residues: 1-855 <SRI>
A:Cross-references: GB:K03458; GB:M16322; NID:g329398; PIDN:AAA45380.1; PID:g329403
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-855/Product: env polyprotein #status predicted <MAT>
F:20-500/Product: exterior membrane glycoprotein #status predicted <EXT>
F:501-855/Product: transmembrane glycoprotein #status predicted <TMM>
F:87,129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,40
Query Match 89.6%; Score 43; DB 1; Length 855;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
Db 668 LWNWFNI 674
|||||:|

Search completed: February 25, 2004, 15:45:39
Job time : 4.25949 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:39 ; Search time 1.41772 Seconds
(without alignments)
257.096 Million cell updates/sec

Title: US-09-877-606-5
Perfect score: 48
Sequence: 1 LWNWFDI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt.42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 48 | 100.0 | 843 | 1 ENV_HV1Y2 | P35961 human immun |
| 2 | 48 | 100.0 | 847 | 1 ENV_HV1S1 | P19550 human immun |
| 3 | 48 | 100.0 | 847 | 1 ENV_HV1W2 | P05880 human immun |
| 4 | 48 | 100.0 | 854 | 1 ENV_SIVCZ | P17281 chimpanzee |
| 5 | 48 | 100.0 | 856 | 1 ENV_HV1MN | P05877 human immun |
| 6 | 48 | 100.0 | 856 | 1 ENV_HV1ZH | P05881 human immun |
| 7 | 48 | 100.0 | 865 | 1 ENV_HV1KH | P04579 human immun |
| 8 | 44 | 91.7 | 1028 | 1 FDVG_HAEIN | P46448 haemophilus |
| 9 | 43 | 89.6 | 851 | 1 ENV_HV1B8 | P04582 human immun |
| 10 | 43 | 89.6 | 852 | 1 ENV_HV1BN | P12488 human immun |
| 11 | 43 | 89.6 | 853 | 1 ENV_HV1MF | P19551 human immun |
| 12 | 43 | 89.6 | 853 | 1 ENV_HV1Z2 | P12487 human immun |
| 13 | 43 | 89.6 | 855 | 1 ENV_HV1Z6 | P04580 human immun |
| 14 | 43 | 89.6 | 856 | 1 ENV_HV1B1 | P03375 human immun |
| 15 | 43 | 89.6 | 856 | 1 ENV_HV1H2 | P04578 human immun |
| 16 | 43 | 89.6 | 856 | 1 ENV_HV1H3 | P04624 human immun |
| 17 | 43 | 89.6 | 856 | 1 ENV_HV1LM | Q70626 human immun |
| 18 | 43 | 89.6 | 856 | 1 ENV_HV1SC | P05878 human immun |
| 19 | 43 | 89.6 | 861 | 1 ENV_HV1BR | P03377 human immun |
| 20 | 42 | 87.5 | 846 | 1 ENV_HV1ND | P18799 human immun |
| 21 | 42 | 87.5 | 852 | 1 ENV_HV1S3 | P19549 human immun |
| 22 | 42 | 87.5 | 853 | 1 ENV_HV1EL | P04581 human immun |
| 23 | 42 | 87.5 | 855 | 1 ENV_HV1A2 | P03378 human immun |
| 24 | 42 | 87.5 | 856 | 1 ENV_HV1W1 | P31872 human immun |
| 25 | 42 | 87.5 | 859 | 1 ENV_HV1MA | P04583 human immun |
| 26 | 42 | 87.5 | 861 | 1 ENV_HV1KB | P31819 human immun |
| 27 | 42 | 87.5 | 863 | 1 ENV_HV1Z8 | P05882 human immun |
| 28 | 41 | 85.4 | 848 | 1 ENV_HV1JR | P20871 human immun |
| 29 | 41 | 85.4 | 867 | 1 ENV_HV1J3 | P12489 human immun |
| 30 | 40 | 83.3 | 606 | 1 ABD4_HUMAN | O14678 homo sapien |
| 31 | 40 | 83.3 | 606 | 1 ABD4_MOUSE | O89016 mus musculu |
| 32 | 38 | 79.2 | 241 | 1 CTR3_YEAST | Q06686 saccharomyc |
| 33 | 38 | 79.2 | 285 | 1 HA33_CLOBO | P46084 clostridium |

34 38 79.2 305 1 BLAI_HAEIN P33949 haemophilus
35 38 79.2 359 1 PEXC_HUMAN O00623 homo sapien
36 38 79.2 359 1 PEXC_MOUSE Q8VC48 mus musculu
37 38 79.2 367 1 POP2_MOUSE Q9EB82 mus musculu
38 38 79.2 483 1 PRPD_HAEIN P44817 haemophilus
39 38 79.2 562 1 SYR2_BACAA Q81R81 bacillus an
40 38 79.2 665 1 CNG_DROME Q24278 drosophila
41 37 77.1 855 1 ENV_HV1OY P20888 human immun
42 37 77.1 856 1 ENV_HV1PV P03376 human immun
43 37 77.1 1616 1 P200_MYCGE Q49429 mycoplasma
44 36 75.0 328 1 GRHR_CAVPO Q8CH60 cavia porce
45 36 75.0 785 1 ISP4_SCHPO P40900 schizosacch

ALIGNMENTS

RESULT 1
ENV_HV1Y2
ID ENV_HV1Y2 STANDARD; PRT; 843 AA.
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
properties of human immunodeficiency virus type 1 in vivo: evidence
for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; M93258; -; NOT_ANNOTATED_CDS.
DR PIR; H44001; H44001.
DR FDB; IG9N; 27-DEC-00.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 490 843 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 738 755 POTENTIAL.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 201 BY SIMILARITY.
FT DISULFID 125 192 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 214 243 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 292 326 BY SIMILARITY.
FT DISULFID 373 432 BY SIMILARITY.
FT DISULFID 380 405 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 381 381 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 803 803 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

 Query Match 100.0%; Score 48; DB 1; Length 843;
 Best Local Similarity 100.0%; Pred. No. 3.7; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

 QY 1 LMNWFPI 7
 Db 656 LMNWFPI 662
 |||||

 RESULT 2
 ENV_HV1S1 STANDARD; PRT; 847 AA.
 AC P19550;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
 OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90347835; PubMed=2384920;
 RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
 RA "Viral determinants of human immunodeficiency virus type 1 T-cell or
 RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
 RL J. Virol. 64:4390-4398(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M55024; AAA45072.1; --
 CC PDB; LOBE; 15-MAY-97.
 CC HIV; M38428; ENVSSFI62.
 CC InterPro; IPR000328; Env_GP41.
 CC Pfam; PF00516; GP120; 1.
 CC Pfam; PF00517; GP41; 1.
 CC AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;

KW 3D-structure. 1 29
 FT SIGNAL 30 502
 FT CHAIN 503 847
 FT DISULFID 53 73
 FT DISULFID 118 203
 FT DISULFID 125 194
 FT DISULFID 130 155
 FT DISULFID 216 245
 FT DISULFID 226 237
 FT DISULFID 294 328
 FT DISULFID 374 435
 FT DISULFID 381 408
 FT CARBOHYD 87 87
 FT CARBOHYD 135 135
 FT CARBOHYD 154 154
 FT CARBOHYD 186 186
 FT CARBOHYD 195 195
 FT CARBOHYD 232 232
 FT CARBOHYD 239 239
 FT CARBOHYD 260 260
 FT CARBOHYD 274 274
 FT CARBOHYD 293 293
 FT CARBOHYD 299 299
 FT CARBOHYD 329 329
 FT CARBOHYD 336 336
 FT CARBOHYD 352 352
 FT CARBOHYD 382 382
 FT CARBOHYD 388 388
 FT CARBOHYD 392 392
 FT CARBOHYD 398 398
 FT CARBOHYD 401 401
 FT CARBOHYD 438 438
 FT CARBOHYD 454 454
 FT CARBOHYD 602 602
 FT CARBOHYD 607 607
 FT CARBOHYD 616 616
 FT CARBOHYD 628 628
 SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

 Query Match 100.0%; Score 48; DB 1; Length 847;
 Best Local Similarity 100.0%; Pred. No. 3.7; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

 QY 1 LMNWFPI 7
 Db 660 LMNWFPI 666
 |||||

 RESULT 3
 ENV_HV1W2 STANDARD; PRT; 847 AA.
 ID ENV_HV1W2
 AC P05880;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
 OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11705;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86235450; PubMed=3012778;
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
 RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
 RL at risk for AIDS.";
 RL Science 232:1548-1553(1986).
 CC -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
 CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
 CC WAS PERINATALLY INFECTED BY HER MOTHER.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; M12507; AAB12990.1; -;
 DR HIV; M12507; ENV5WJ2;
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT CHAIN 1 29
 FT CHAIN 502 847 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 118 202 BY SIMILARITY.
 FT DISULFID 125 193 BY SIMILARITY.
 FT DISULFID 130 152 BY SIMILARITY.
 FT DISULFID 215 244 BY SIMILARITY.
 FT DISULFID 225 236 BY SIMILARITY.
 FT DISULFID 293 326 BY SIMILARITY.
 FT DISULFID 372 435 BY SIMILARITY.
 FT DISULFID 379 408 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 847 AA; 96466 MW; CD1E33D73AA5BCAE CRC64;

Query Match 100.0%; Score 48; DB 1; Length 847;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 660 LWNWFDI 666
 |||||

RESULT 4
 ENV_SIVCZ STANDARD; PRT; 854 AA.
 AC F17281;
 DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 DE ENV.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90259077; PubMed=2188136;
 RA Huot T., Cheynier R., Meyerhans A., Roelants G., Main-Hobson S.;
 RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";
 RL Nature 345:356-359(1990).
 CC 1- SIMILARITY: STRONG, WITH HIV-1 ENV POLYPROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52154; CAA36407.1; -;
 DR PIR; S09990; VCLJ51.
 DR HIV; X52154; ENVSCPZ.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT CHAIN 1 30
 FT CHAIN 31 500 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT TRANSMEM 501 854 TRANSMEMBRANE GLYCOPROTEIN.
 FT TRANSMEM 517 517 POTENTIAL.
 FT TRANSMEM 675 693 POTENTIAL.
 FT TRANSMEM 805 821 POTENTIAL.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 311 311 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 854 AA; 95803 MW; 2E249AFAD4F2D9B3 CRC64;

Query Match 100.0%; Score 48; DB 1; Length 854;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 |||||

FT DISULFID 130 153 BY SIMILARITY. (POTENTIAL).
 FT DISULFID 219 248 BY SIMILARITY. (POTENTIAL).
 FT DISULFID 229 240 BY SIMILARITY. (POTENTIAL).
 FT DISULFID 297 331 BY SIMILARITY. (POTENTIAL).
 FT DISULFID 379 445 BY SIMILARITY. (POTENTIAL).
 FT DISULFID 386 418 BY SIMILARITY. (POTENTIAL).
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 95909 MW; 839653P8BBD174E CRC64;
 Query Match 100.0%; Score 48; DB 1; Length 856;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWNWFDI 7
 |||||
 Db 669 LWNWFDI 675
 RESULT 7
 ENV_HVIRH STANDARD; PRT; 865 AA.
 AC P04579;
 DT 13-AUG-1987 (Rel. 05, Last Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11701;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86218077; PubMed=2423250;
 RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
 RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal P.;
 RT "Identification and characterization of conserved and variable
 regions in the envelope gene of HTLV-III/LAV, the retrovirus of
 AIDS.";
 RT AIDS.;
 RL Cell 45:637-648(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M17451; AAA45057.1; -;
 DR HIV; M17451; ENVSRF.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 519 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 520 865 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 218 BY SIMILARITY.
 FT DISULFID 125 209 BY SIMILARITY.
 FT DISULFID 130 157 BY SIMILARITY.
 FT DISULFID 231 260 BY SIMILARITY.
 FT DISULFID 241 252 BY SIMILARITY.
 FT DISULFID 309 343 BY SIMILARITY.
 FT DISULFID 389 452 BY SIMILARITY.
 FT DISULFID 396 425 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 254 254 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 367 367 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 825 825 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 865 AA; 97809 MW; 28928BC2314ADCAC CRC64;
 Query Match 100.0%; Score 48; DB 1; Length 865;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWNWFDI 7
 |||||
 Db 678 LWNWFDI 684
 RESULT 8
 FDHG_HAEIN STANDARD; PRT; 1028 AA.
 ID FDHG_HAEIN
 AC P46448;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase major subunit (EC 1.2.1.2) (Formate
DE dehydrogenase alpha subunit) (PDB alpha subunit).
GN PDXG OR HI0006
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Rd / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Klenavanne K., Suttan G., Fitchugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Keiley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brand R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -I- FUNCTION: ALLOWS TO USE FORMATE AS MAJOR ELECTRON DONOR DURING
CC ANAEROBIC RESPIRATION. SUBUNIT ALPHA POSSIBLY FORMS THE ACTIVE
CC SITE.
CC -I- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -I- COFACTOR: Molybdenum (molybdopterin) and selenocysteine. The
CC active-site selenocysteine is encoded by the opal codon, UGA. May
CC bind a 4Fe-4S cluster.
CC -I- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
CC BY SUBUNITS ALPHA, BETA AND GAMMA.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -I- SIMILARITY: ORTHOLOG OF BOTH E.COLI FDNG AND FDGO.
CC -I- SIMILARITY: Belongs to the prokaryotic molybdopterin-containing
CC oxidoreductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32686; -; NOT_ANNOTATED_CDS.
DR TIGR; HI0006; -;
DR InterPro; IPR009010; Asp decarb fold.
DR InterPro; IPR006443; Formate-dh-alpha.
DR InterPro; IPR006657; Mol dinuc bind.
DR InterPro; IPR006963; Molybdop_Fe4S4.
DR InterPro; IPR006656; Molybdop_Fe4S4.
DR InterPro; IPR006655; Molybdop_Fe4S4.
DR InterPro; IPR006311; Tat.
DR Pfam; PF04879; Molybdop_Fe4S4; 1.
DR Pfam; PF00384; molybdop_Fe4S4; 1.
DR Pfam; PF01568; Molybdop_binding; 1.
DR TIGRFAMs; TIGR01553; formate-DH-alpha; 1.
DR TIGRFAMs; TIGR01409; Tat signal seq; 1.
DR PROSITE; PS00551; Molybdop_Fe4S4; 1.
DR PROSITE; PS00430; Molybdop_Fe4S4; 1.
DR PROSITE; PS00932; Molybdop_Fe4S4; 1.
DR Oxidoreductase; Molybdenum; Selenocysteine; Selenium; NAD;
KW Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 50 50 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 53 53 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 57 57 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 100 100 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT SE_CYS 204 204 PROBABLE.
SQ SEQUENCE 1028 AA; 115403 MW; 295F185BB020EDF9 CRC64;
Query Match 91.7%; Score 44; DB 1; Length 1028;

Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 WNWFDI 7
DB 812 WNWFDI 817
[1]
RESULT 9
ENV_HV1B8 STANDARD; PRT; 851 AA.
ID ENV_HV1B8
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GPI60 precursor [Contains: Exterior membrane
DE glycoprotein (GPI20); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroviridae; Retrovirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R., Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02011; AAA44661.1; -;
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HHG; 31-OCT-93.
DR PDB; 1QQ3; 02-JAN-00.
DR PDB; 1SZT; 24-DEC-97.
DR HIV; K02011; ENV5B8.
DR GlycoSuiteDB; P04582; -;
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 31 506 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 507 851 BY SIMILARITY.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 440 BY SIMILARITY.
FT DISULFID 385 413 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 851;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMNWFPI 7
|||||:
Db 664 LMNWFNI 670

RESULT 10
ENV HV1BN STANDARD; PRT; 852 AA.
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 10-OCT-1989 (Rel. 12, Last sequence update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BR isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=11693;
RX SEQUENCE FROM N.A.
RA MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA Dandekar S.;
RA "Biological and molecular characterization of human immunodeficiency
virus (HIV-1BR) from the brain of a patient with progressive
dementia.";
RL Virology 168:79-89 (1989).
CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21098; AAA44221.1; -.
CC PIR; A31667; VCLJBR.
CC PDB; 1IM7; 23-OCT-02.
CC HIV; M21098; ENV$BRVA.
CC InterPro; IPR000328; Env GP41.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
```

```
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure. 1 30
FT SIGNAL 31 507 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 508 852 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 155 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 439 BY SIMILARITY.
FT DISULFID 383 412 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match 89.6%; Score 43; DB 1; Length 852;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMNWFPI 7
|||||:
Db 665 LMNWFNI 671

RESULT 11
ENV HV1MF STANDARD; PRT; 853 AA.
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
ON NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamonic C., Mann A.M., Meier C.,
RA Wasiaik A.;
RA "Cloning and characterization of human immunodeficiency virus type 1
```

RT variants diminished in the ability to induce syncytium-independent
 RT cytolysis".
 RL J. Virol. 64:3792-3803 (1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M33943; AAA44850.1; --
 DR PDB; 1A1K; 16-JUN-97.
 DR HIV; M33943; ENV\$MFA.
 DR InterPro; IPR000328; Env GP41.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
 KW 3D-structure. 1 30
 FT SIGNAL 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 510 853 TRANSMEMBRANE GLYCOPROTEIN.
 FT CHAIN 54 74 BY SIMILARITY.
 FT DISULFID 119 203 BY SIMILARITY.
 FT DISULFID 126 194 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 216 245 BY SIMILARITY.
 FT DISULFID 226 237 BY SIMILARITY.
 FT DISULFID 294 329 BY SIMILARITY.
 FT DISULFID 376 443 BY SIMILARITY.
 FT DISULFID 383 416 BY SIMILARITY.
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 672 672 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 814 814 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 853 AA; 96912 MW; 33778993B6F22ABA CRC64;
 Query Match 89.6%; Score 43; DB 1; Length 853;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LMNWFPI 7
 |||||:
 Db 667 LMNWFPI 673

RESULT 12
 ENV_HVIZ2
 ID ENV_HVIZ2 STANDARD; PRT; 853 AA.
 AC P12487;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (Z2/CDC-Z34 isolate) (HIV-1).
 OS Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11683;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Theodore T., Buckler-White A.;
 RL Submitted (NOV-1988) to the HIV data bank.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22639; AAA45370.1; --
 DR PIR; S54384; S54384.
 DR HIV; M22639; ENV\$Z226.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 31
 FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 206 BY SIMILARITY.
 FT DISULFID 125 197 BY SIMILARITY.
 FT DISULFID 130 154 BY SIMILARITY.
 FT DISULFID 219 248 BY SIMILARITY.
 FT DISULFID 229 240 BY SIMILARITY.
 FT DISULFID 287 330 BY SIMILARITY.
 FT DISULFID 376 442 BY SIMILARITY.
 FT DISULFID 383 415 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).

```

FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B0B8CBAPF7008 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 853;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNNWFDI 7
Db 666 LNNWFNI 672

RESULT 13
ENV_HV126
ID ENV_HV126 STANDARD; PRT; 855 AA.
AC P04580;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K03458; AAA5380.1; -.
DR PIR; D26192; VCLJZR.
DR HIV; K03458; ENV$Z6.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT CHAIN 1 30
FT CHAIN 31 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 511 855 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 207 BY SIMILARITY.
FT DISULFID 125 198 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 220 249 BY SIMILARITY.
FT DISULFID 230 241 BY SIMILARITY.
FT DISULFID 298 332 BY SIMILARITY.
FT DISULFID 378 444 BY SIMILARITY.
FT DISULFID 385 417 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 855 AA; 96971 MW; 3B4D3D6E239C3457 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 855;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNNWFDI 7
Db 668 LNNWFNI 674

RESULT 14
ENV_HV181
ID ENV_HV181 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
RT potential glycosylation sites of the type 1 recombinant human
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in
RT Chinese hamster ovary cells.";
RL J. Biol. Chem. 265:10373-10382(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M15654; AAA44205.1; -;
 DR PIR; A03973; VCLJH3.
 DR HIV; M15654; ENV5BH102.
 DR InterPro; IPR000328; Env GP41.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97224 MW; 0BFPB1A18931BB27 CRC64;
 Query Match 89.6%; Score 43; DB 1; Length 856;
 Best Local Similarity 85.9%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LNNWFDI 7
 |||||:
 Db 669 LNNWFNI 675
 RESULT 15
 ENV_HV1H2
 ID ENV_HV1H2 STANDARD; PRT; 856 AA.

AC P04578; O09779;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OK NCBI_TaxID=11706;
 RN [1]_TaxID=11706;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87299196; PubMed=3040055;
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
 RA Gallo R.C., Wong-Staal F.;
 RT "Complete nucleotide sequences of functional clones of the AIDS
 RT virus";
 RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
 RN [2]
 RP REVISIONS.
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
 RA Gallo R.C., Wong-Staal F.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; K03455; AAB50362.1; -;
 DR EMBL; AP038399; AAB99976.1; -;
 DR PDB; 1DF4; 26-JAN-00.
 DR PDB; 1DF5; 26-JAN-00.
 DR PDB; 1DLB; 02-APR-00.
 DR PDB; 1G9M; 27-DEC-00.
 DR PDB; 1GCL; 19-AUG-98.
 DR PDB; 1GZL; 10-OCT-02.
 DR PDB; 1K33; 10-OCT-01.
 DR PDB; 1K34; 10-OCT-01.
 DR HIV; K03455; ENV5HXB2.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
 KW 3D-structure.
 FT SIGNAL 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 31 511 TRANSMEMBRANE GLYCOPROTEIN.
 FT CHAIN 512 856 BY SIMILARITY.
 FT DISULFID 54 74 BY SIMILARITY.
 FT DISULFID 119 205 BY SIMILARITY.
 FT DISULFID 126 196 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 331 BY SIMILARITY.
 FT DISULFID 378 445 BY SIMILARITY.
 FT DISULFID 385 418 BY SIMILARITY.
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).

| | | | | |
|----|----------|---|------------------------|--------------|
| FT | CARBOHYD | 276 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 289 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 295 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 301 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 332 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 339 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 356 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 386 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 392 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 397 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 406 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 448 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 463 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 611 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 616 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 624 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 637 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 674 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 750 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 816 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| SQ | SEQUENCE | 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64; | | |

Query Match 89.6%; Score 43; DB 1; Length 856;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 |||||:
 Db 669 LWNWFNI 675

Search completed: February 25, 2004, 15:41:47
 Job time : 3.41772 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:35:54 ; Search time 6.86709 Seconds
(without alignments)
321.625 Million cell updates/sec

Title: US-09-877-606-5

Perfect score: 48

Sequence: 1 LWNWFDI 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_25.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_human.*

6: sp_invertebrate.*

7: sp_mammal.*

8: sp_mhc.*

9: sp_organelle.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 48 | 100.0 | 41 | 15 | Q69891 human immun |
| 2 | 48 | 100.0 | 41 | 15 | Q69892 human immun |
| 3 | 48 | 100.0 | 117 | 15 | Q9QON5 human immun |
| 4 | 48 | 100.0 | 117 | 15 | Q9YRT2 human immun |
| 5 | 48 | 100.0 | 117 | 15 | Q9YRS9 human immun |
| 6 | 48 | 100.0 | 117 | 15 | Q9YRS0 human immun |
| 7 | 48 | 100.0 | 117 | 15 | Q9YRT7 human immun |
| 8 | 48 | 100.0 | 117 | 15 | Q9YRS1 human immun |
| 9 | 48 | 100.0 | 117 | 15 | Q9WKU0 human immun |
| 10 | 48 | 100.0 | 117 | 15 | Q9YRT6 human immun |
| 11 | 48 | 100.0 | 117 | 15 | Q9YRR7 human immun |
| 12 | 48 | 100.0 | 117 | 15 | Q9YRU0 human immun |
| 13 | 48 | 100.0 | 117 | 15 | Q7SVI8 human immun |
| 14 | 48 | 100.0 | 117 | 15 | Q7SVIS human immun |
| 15 | 48 | 100.0 | 117 | 15 | Q7SVI3 human immun |
| 16 | 48 | 100.0 | 117 | 15 | Q7SVI2 human immun |

| | | | | | |
|----|----|-------|-----|----|--------------------|
| 17 | 48 | 100.0 | 117 | 15 | Q7SVI0 human immun |
| 18 | 48 | 100.0 | 117 | 15 | Q7SVH9 human immun |
| 19 | 48 | 100.0 | 117 | 15 | Q7SVH8 human immun |
| 20 | 48 | 100.0 | 117 | 15 | Q7SVH6 human immun |
| 21 | 48 | 100.0 | 117 | 15 | Q7SVH5 human immun |
| 22 | 48 | 100.0 | 117 | 15 | Q7SVH4 human immun |
| 23 | 48 | 100.0 | 117 | 15 | Q7SVH3 human immun |
| 24 | 48 | 100.0 | 117 | 15 | Q7SVH2 human immun |
| 25 | 48 | 100.0 | 117 | 15 | Q7SVH0 human immun |
| 26 | 48 | 100.0 | 117 | 15 | Q7SVG9 human immun |
| 27 | 48 | 100.0 | 117 | 15 | Q7SVG8 human immun |
| 28 | 48 | 100.0 | 117 | 15 | Q7SVG7 human immun |
| 29 | 48 | 100.0 | 117 | 15 | Q7SVG6 human immun |
| 30 | 48 | 100.0 | 117 | 15 | Q7SVG5 human immun |
| 31 | 48 | 100.0 | 117 | 15 | Q7SVG4 human immun |
| 32 | 48 | 100.0 | 117 | 15 | Q7SVG3 human immun |
| 33 | 48 | 100.0 | 117 | 15 | Q7SVF9 human immun |
| 34 | 48 | 100.0 | 117 | 15 | Q7SVF7 human immun |
| 35 | 48 | 100.0 | 117 | 15 | Q7SVF6 human immun |
| 36 | 48 | 100.0 | 117 | 15 | Q7SVF4 human immun |
| 37 | 48 | 100.0 | 117 | 15 | Q7SVF2 human immun |
| 38 | 48 | 100.0 | 117 | 15 | Q7SVF1 human immun |
| 39 | 48 | 100.0 | 117 | 15 | Q7SVE9 human immun |
| 40 | 48 | 100.0 | 117 | 15 | Q7SVE7 human immun |
| 41 | 48 | 100.0 | 117 | 15 | Q7SVE6 human immun |
| 42 | 48 | 100.0 | 117 | 15 | Q7SVE5 human immun |
| 43 | 48 | 100.0 | 117 | 15 | Q7SV48 human immun |
| 44 | 48 | 100.0 | 118 | 15 | Q9E5R1 human immun |
| 45 | 48 | 100.0 | 118 | 15 | Q9E5P8 human immun |

ALIGNMENTS

RESULT 1

Q69891 ID Q69891 PRELIMINARY; PRT; 41 AA.
AC Q69891;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp41 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94211861; PubMed=7512731;
RA Conley A.J., Kessler J.A., Boots L.J., Tung J.S., Arnold B.A.,
RA Keller P.M., Shaw A.R., Emini E.A.;
RT "Neutralization of divergent human immunodeficiency virus type 1
RT variants and primary isolates by 1AM-41-2F5, an anti-gp41 human
RT monoclonal antibody.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3348-3352(1994).
DR EMBL; U06721; AAA19134.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
KW Pfam; PF00517; GP41; 1.
FT Transmembrane.
FT NON_TER 1 1
FT NON_TER 41 41
SQ SEQUENCE 41 AA; 5082 MW; 089C7E2CDD403CC6 CRC64;
Query Match 100.0%; Score 48; DB 15; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.79; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
QY 1 LWNWFDI 7
DB 30 LWNWFDI 36

RESULT 2
 Q69892 ID Q69892 PRELIMINARY; PRT; 41 AA.
 AC Q69892;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Gp41 (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94211861; PubMed=7512731;
 RA Conley A.J., Keesler J.A., Boots L.J., Tung J.S., Arnold B.A.,
 RA Keller P.M., Shaw A.R., Emini E.A.;
 RT "Neutralization of divergent human immunodeficiency virus type 1
 RT variants and primary isolates by IAW-41-2F5, an anti-gp41 human
 RT monoclonal antibody.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3348-3352(1994).
 DR EMBL; U06722; AAA19135.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.
 KW Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 41 41
 SQ SEQUENCE 41 AA; 5081 MW; 0B9C70CC33403CC6 CRC64;

 Query Match 100.0%; Score 48; DB 15; Length 41;
 Best Local Similarity 100.0%; Pred. No. 0.79;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 30 LWNWFDI 36

RESULT 3
 Q9QQN5 ID Q9QQN5 PRELIMINARY; PRT; 117 AA.
 AC Q9QQN5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein gp41 (fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=96USHIPS7;
 RA Ellenberger D., Sullivan P.S., Dorn J., Schable C., Spira T.J.,
 RA Folks T.M., Lal R.B.;
 RT "Viral and Immunologic Examination of HIV-1-infected Persistently
 RT Seronegative (HPS) Persons.";
 RL J. Infect. Dis. 0:0-0(1999).
 DR EMBL; AF157468; AAD45888.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.
 KW Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 14011 MW; 9A71687C21470B60 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 109 LWNWFDI 115

RESULT 4
 Q9VRT2 ID Q9VRT2 PRELIMINARY; PRT; 117 AA.
 AC Q9VRT2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=96USNG15;
 RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
 RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
 RA Jones J., Ward J.;
 RT "Surveillance of Central African Nationals living in the United States
 RT reveals multiple subtypes of HIV-1 Group M and Group O.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF096334; AAD04409.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.
 KW Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 13959 MW; DED23BB002B524D9 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFDI 7
 DB 109 LWNWFDI 115

RESULT 5
 Q9YRS9 ID Q9YRS9 PRELIMINARY; PRT; 117 AA.
 AC Q9YRS9;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=96USSN94;
 RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
 RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
 RA Jones J., Ward J.;
 RT "Surveillance of Central African Nationals living in the United States
 RT reveals multiple subtypes of HIV-1 Group M and Group O.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF096337; AAD04412.1; -.

```

DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13953 MW; F1A9174D1A50B77B CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
DB 109 LNNWFDI 115

RESULT 6
QYRSO PRELIMINARY; PRT; 117 AA.
AC QYRSO;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=97USCN31;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.;
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096346; AAD04421.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14006 MW; 7440C1BA73F95719 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
DB 109 LNNWFDI 115

RESULT 7
QYRT7 PRELIMINARY; PRT; 117 AA.
AC QYRT7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=96USNG17;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.;
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096328; AAD04403.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13985 MW; 959F54B7B737E9E CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
DB 109 LNNWFDI 115

RESULT 8
QYRS1 PRELIMINARY; PRT; 117 AA.
AC QYRS1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=96USCF72;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.;
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096345; AAD04420.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13925 MW; 09DF32E239BD59D0 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFDI 7
DB 109 LNNWFDI 115

RESULT 9
Q9WKU0 PRELIMINARY; PRT; 117 AA.
AC Q9WKU0;

```

DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RC STRAIN=96USNG19;
 RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
 RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
 RA Jones J., Ward J.;
 RA "Surveillance of Central African Nationals living in the United States
 RT reveals multiple subtypes of HIV-1 Group M and Group O.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF096333; AAD04408.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.
 KW Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 13959 MW; BBA0CA6FDE179EBC CRC64;
 Query Match 100.0%; Score 48; DB 15; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LMNWFPI 7
 Db 109 LMNWFPI 115
 RESULT 10
 QYRT6 PRELIMINARY; PRT; 117 AA.
 AC QYRT6;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RC STRAIN=96USNG58;
 RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
 RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
 RA Jones J., Ward J.;
 RA "Surveillance of Central African Nationals living in the United States
 RT reveals multiple subtypes of HIV-1 Group M and Group O.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF096329; AAD04404.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.
 KW Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 13962 MW; 83A17E3F6154966C CRC64;
 Query Match 100.0%; Score 48; DB 15; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFPI 7
 Db 109 LMNWFPI 115
 RESULT 11
 QYRR7 PRELIMINARY; PRT; 117 AA.
 AC QYRR7;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RC STRAIN=97USNG30;
 RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
 RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
 RA Jones J., Ward J.;
 RA "Surveillance of Central African Nationals living in the United States
 RT reveals multiple subtypes of HIV-1 Group M and Group O.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF096349; AAD04424.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.
 KW Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 14003 MW; 8C093B3E935BCEP CRC64;
 Query Match 100.0%; Score 48; DB 15; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LMNWFPI 7
 Db 109 LMNWFPI 115
 RESULT 12
 QYRU0 PRELIMINARY; PRT; 117 AA.
 AC QYRU0;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RC STRAIN=96USCF70;
 RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
 RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
 RA Jones J., Ward J.;
 RA "Surveillance of Central African Nationals living in the United States
 RT reveals multiple subtypes of HIV-1 Group M and Group O.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF096325; AAD04400.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000328; Env_GP41.
 DR Pfam; PF00517; GP41; 1.

```

KW Transmembrane. 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13941 MW; F26E40AD3DFB1B60 CRC64;
  Query Match 100.0%; Score 48; DB 15; Length 117;
  Best Local Similarity 100.0%; Pred. No. 2.2;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
  |||||
Db 109 LMNWFDI 115

RESULT 13
Q7SVI8 PRELIMINARY; PRT; 117 AA.
AC Q7SVI8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein gp41 domain (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99NGG681;
RA Zeh C., Agwale S.M., Robbins K.E., Odama L., Sani-Gwarco N.,
RA Gboun M.S., Pieniazek D., Wambele C., Folks T.M., Kalish M.L.;
RT "Presence of HIV-2 subtype B in single and HIV-2 subtype A in dual
RT HIV-1/HIV-2 infections in Nigeria.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV102796; AAM51891.1; -
KW Envelope protein.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14076 MW; F9B2B5A3C415C6F6 CRC64;
  Query Match 100.0%; Score 48; DB 15; Length 117;
  Best Local Similarity 100.0%; Pred. No. 2.2;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
  |||||
Db 109 LMNWFDI 115

RESULT 14
Q7SVI5 PRELIMINARY; PRT; 117 AA.
AC Q7SVI5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein gp41 domain (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99NGI063;
RA Zeh C., Agwale S.M., Robbins K.E., Odama L., Sani-Gwarco N.,
RA Gboun M.S., Pieniazek D., Wambele C., Folks T.M., Kalish M.L.;
RT "Presence of HIV-2 subtype B in single and HIV-2 subtype A in dual
RT HIV-1/HIV-2 infections in Nigeria.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV102799; AAM51894.1; -
KW Envelope protein.
FT NON_TER 1
FT NON_TER 117

```

```

SQ SEQUENCE 117 AA; 13951 MW; 3B7A7B3ABDA164C0 CRC64;
  Query Match 100.0%; Score 48; DB 15; Length 117;
  Best Local Similarity 100.0%; Pred. No. 2.2;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
  |||||
Db 109 LMNWFDI 115

RESULT 15
Q7SVI3 PRELIMINARY; PRT; 117 AA.
AC Q7SVI3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein gp41 domain (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99NGI075;
RA Zeh C., Agwale S.M., Robbins K.E., Odama L., Sani-Gwarco N.,
RA Gboun M.S., Pieniazek D., Wambele C., Folks T.M., Kalish M.L.;
RT "Presence of HIV-2 subtype B in single and HIV-2 subtype A in dual
RT HIV-1/HIV-2 infections in Nigeria.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV102801; AAM51896.1; -
KW Envelope protein.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13988 MW; 6AA617B7C19E7EEB CRC64;
  Query Match 100.0%; Score 48; DB 15; Length 117;
  Best Local Similarity 100.0%; Pred. No. 2.2;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFDI 7
  |||||
Db 109 LMNWFDI 115

Search completed: February 25, 2004, 15:44:39
Job time : 8.86709 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:09 ; Search time 10.3228 Seconds
(without alignments)
191.599 Million cell updates/sec

Title: US-09-877-606-7

Perfect score: 48

Sequence: 1 LMNWFNI 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 48 | 100.0 | 15 | 1 AAP90238 | Aap90238 Antigenic |
| 2 | 48 | 100.0 | 15 | 4 AAU12542 | Aau12542 Human HIV |
| 3 | 48 | 100.0 | 15 | 4 AAU12532 | Aau12532 Human HIV |
| 4 | 48 | 100.0 | 15 | 4 AAU12524 | Aau12524 Human HIV |
| 5 | 48 | 100.0 | 20 | 2 AAU07920 | Aau07920 GP41 pept |
| 6 | 48 | 100.0 | 23 | 1 AAP82469 | Aap82469 Peptide c |
| 7 | 48 | 100.0 | 23 | 2 AAU07974 | Aau07974 HIV pepti |
| 8 | 48 | 100.0 | 23 | 4 AAU12483 | Aau12483 HIV Th-CT |
| 9 | 48 | 100.0 | 24 | 4 AAU12488 | Aau12488 HIV Th-CT |
| 10 | 48 | 100.0 | 24 | 5 AAU70251 | Aau70251 HIV Th-CT |
| 11 | 48 | 100.0 | 32 | 4 AAU12479 | Aau12479 HIV Th-CT |
| 12 | 48 | 100.0 | 32 | 5 AAU70242 | Aau70242 HIV Th-CT |
| 13 | 48 | 100.0 | 33 | 5 ABB09199 | Abb09199 HIV-1 gp1 |
| 14 | 48 | 100.0 | 36 | 2 AAU17022 | Aau17022 DP-178-de |
| 15 | 48 | 100.0 | 36 | 3 AAY89191 | Aay89191 Core poly |
| 16 | 48 | 100.0 | 36 | 3 AAY89199 | Aay89199 Core poly |
| 17 | 48 | 100.0 | 36 | 3 AAY89188 | Aay89188 Core poly |
| 18 | 48 | 100.0 | 36 | 3 AAY89196 | Aay89196 Core poly |
| 19 | 48 | 100.0 | 36 | 3 AAY89197 | Aay89197 Core poly |
| 20 | 48 | 100.0 | 36 | 3 AAY89189 | Aay89189 Core poly |
| 21 | 48 | 100.0 | 36 | 3 AAY89194 | Aay89194 Core poly |
| 22 | 48 | 100.0 | 36 | 3 AAY89198 | Aay89198 Core poly |
| 23 | 48 | 100.0 | 36 | 3 AAY89200 | Aay89200 Core poly |
| 24 | 48 | 100.0 | 36 | 3 AAY89190 | Aay89190 Core poly |
| 25 | 48 | 100.0 | 36 | 3 AAY89193 | Aay89193 Core poly |

ALIGNMENTS

RESULT 1

AAP90238
ID AAP90238 standard; peptide; 15 AA.

XX AC

XX AAP90238;

XX AC

XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 26-FEB-1990 (first entry)

XX XX

DE Antigenic peptide for detecting, inhibiting and neutralising HIV-1.

XX XX

XX HIV-1; antigenic determinants.

XX XX

OS Human immunodeficiency virus 1.

XX XX

PN EP330359-A.

XX XX

PD 30-AUG-1989.

XX XX

PF 14-FEB-1989; 89EP-00301364.

XX XX

PR 25-FEB-1988; 88US-00160378.

XX XX

PA (BIRA) BIO RAD LAB INC.

XX XX

PI Walker RP, Parekh BS;

XX XX

DR WPI; 1989-250452/35.

XX XX

PT New peptide(s) for detecting, inhibiting and neutralising HIV-1 -
corresp. to antigenic determinants encoded by conserved regions of HIV-1
genome.

XX XX

PS Disclosure; Page 4; 16pp; English.

XX XX

CC Peptide is one of several fragments from gp 120 and gp41 (env gene) p18
and p24 (gag gene), p32 (pol gene) and proteins encoded by the tat, orf
trg/art and sor genes. They are used for detecting, inhibiting and
neutralising HIV-1 infection. Dosage is pref. 5-25 mg/kg. The peptides
can be used for any type of immunological detection esp. dot blot and
ELISA. See also AAP9191-P90274. (Updated on 25-MAR-2003 to correct PA
field.) (Updated on 24-OCT-2003 to standardise OS field)

XX XX

SQ Sequence 15 AA;

XX XX

Query Match 100.0%; Score 48; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.66;

Aay89195 Core poly
Aab77542 Core poly
Aab77547 Core poly
Aab77549 Core poly
Aab77552 Core poly
Aab77553 Core poly
Aab77543 Core poly
Aab77544 Core poly
Aab77545 Core poly
Aab77546 Core poly
Aab77548 Core poly
Aab77550 Core poly
Aab77551 Core poly
Abb00548 Viral DP1
Abb00552 Viral DP1
Abb02024 Viral cor
Abb00555 Viral DP1
Abb02032 Viral cor
Abb00547 Viral DP1
Abb00549 Viral DP1

26 48 100.0 36 3 AAY89195
27 48 100.0 36 4 AAB77542
28 48 100.0 36 4 AAB77547
29 48 100.0 36 4 AAB77549
30 48 100.0 36 4 AAB77552
31 48 100.0 36 4 AAB77553
32 48 100.0 36 4 AAB77543
33 48 100.0 36 4 AAB77544
34 48 100.0 36 4 AAB77545
35 48 100.0 36 4 AAB77546
36 48 100.0 36 4 AAB77548
37 48 100.0 36 4 AAB77550
38 48 100.0 36 4 AAB77551
39 48 100.0 36 4 ABB00548
40 48 100.0 36 4 ABB00552
41 48 100.0 36 4 ABB02024
42 48 100.0 36 4 ABB00555
43 48 100.0 36 4 ABB02032
44 48 100.0 36 4 ABB00547
45 48 100.0 36 4 ABB00549

```

Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  LWNWFNI 7
      |||||
DB      9  LWNWFNI 15

RESULT 2
AAU12542
ID  AAU12542 standard; peptide; 15 AA.
XX
AC  AAU12542;
XX
KW  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
DT  11-SEP-2003 (revised)
DT  27-SEP-2001 (first entry)
XX
XX  Human HIV-1 Th-CTL overlapping epitope #49.
XX
DE  Human immunodeficiency virus 1.
XX
KW  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
XX  Homo sapiens.
OS
OS  Human immunodeficiency virus 1.
XX
PN  WO200156355-A2.
XX
PD  09-AUG-2001.
XX
PF  05-FEB-2001; 2001WO-US003540.
XX
PR  04-FEB-2000; 2000US-00497497.
XX
PA  (UYDU-) UNIV DUKE.
XX
PI  Haynes BF, Liao H;
XX
DR  WPI; 2001-488827/53.
XX
PT  Vaccine for immunizing against human immunodeficiency virus has mixture
PT  or linear array of peptides comprising immunodominant T-helper epitopes
PT  and major histocompatibility complex cytotoxic T-lymphocyte epitopes.
XX
PS  Disclosure; Page 27; 33pp; English.
XX
CC  The present invention relates to human immunodeficiency virus (HIV) and
CC  in particular to a human leukocyte antigen (HLA)-based HIV vaccine. The
CC  vaccine comprises a mixture or linear array of peptides, or its variants,
CC  where the peptides contain immunodominant T-helper (Th) epitopes and
CC  major histocompatibility complex (MHC) cytotoxic T- lymphocyte (CTL)
CC  epitopes and the linear array of peptides are preferably expressed in
CC  modified Vaccinia ankara. The vaccine is useful for immunising a patient
CC  against HIV and focuses immune response on many dominant and subdominant
CC  CTL epitopes of HIV. DNA or live vectors with linear arrays of CTL
CC  epitopes can be used as either primes or boosts of peptides or of each
CC  other to optimally give CTL anti-HIV responses. The vaccine induces
CC  salutary anti-HIV immune responses. AAU12447-AAU12558 represent the amino
CC  acid sequences of the Th-CTL epitopes and HIV immunogenic peptides used
CC  in the invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ  Sequence 15 AA;

Query Match      100.0%; Score 48; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  LWNWFNI 7
      |||||
DB      3  LWNWFNI 9

RESULT 3
AAU12542
ID  AAU12542 standard; peptide; 15 AA.
XX
AC  AAU12542;
XX
KW  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
DT  11-SEP-2003 (revised)
DT  27-SEP-2001 (first entry)
XX
XX  Human HIV-1 Th-CTL overlapping epitope #49.
XX
DE  Human immunodeficiency virus 1.
XX
KW  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
XX  Homo sapiens.
OS
OS  Human immunodeficiency virus 1.
XX
PN  WO200156355-A2.
XX
PD  09-AUG-2001.
XX
PF  05-FEB-2001; 2001WO-US003540.
XX
PR  04-FEB-2000; 2000US-00497497.
XX
PA  (UYDU-) UNIV DUKE.
XX
PI  Haynes BF, Liao H;
XX
DR  WPI; 2001-488827/53.
XX
PT  Vaccine for immunizing against human immunodeficiency virus has mixture
PT  or linear array of peptides comprising immunodominant T-helper epitopes
PT  and major histocompatibility complex cytotoxic T-lymphocyte epitopes.
XX
PS  Disclosure; Page 27; 33pp; English.
XX
CC  The present invention relates to human immunodeficiency virus (HIV) and
CC  in particular to a human leukocyte antigen (HLA)-based HIV vaccine. The
CC  vaccine comprises a mixture or linear array of peptides, or its variants,
CC  where the peptides contain immunodominant T-helper (Th) epitopes and
CC  major histocompatibility complex (MHC) cytotoxic T- lymphocyte (CTL)
CC  epitopes and the linear array of peptides are preferably expressed in
CC  modified Vaccinia ankara. The vaccine is useful for immunising a patient
CC  against HIV and focuses immune response on many dominant and subdominant
CC  CTL epitopes of HIV. DNA or live vectors with linear arrays of CTL
CC  epitopes can be used as either primes or boosts of peptides or of each
CC  other to optimally give CTL anti-HIV responses. The vaccine induces
CC  salutary anti-HIV immune responses. AAU12447-AAU12558 represent the amino
CC  acid sequences of the Th-CTL epitopes and HIV immunogenic peptides used
CC  in the invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ  Sequence 15 AA;

Query Match      100.0%; Score 48; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  LWNWFNI 7
      |||||
DB      3  LWNWFNI 9

RESULT 4
AAU12524
ID  AAU12524 standard; peptide; 15 AA.
XX
AC  AAU12524;
XX
KW  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
DT  11-SEP-2003 (revised)
DT  27-SEP-2001 (first entry)
XX
XX

```

```

AAU12532
ID  AAU12532 standard; peptide; 15 AA.
XX
AC  AAU12532;
XX
DT  11-SEP-2003 (revised)
DT  27-SEP-2001 (first entry)
XX
XX  Human HIV-1 Th-CTL overlapping epitope #39.
XX
DE  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
XX  Homo sapiens.
OS
OS  Human immunodeficiency virus 1.
XX
PN  WO200156355-A2.
XX
PD  09-AUG-2001.
XX
PF  05-FEB-2001; 2001WO-US003540.
XX
PR  04-FEB-2000; 2000US-00497497.
XX
PA  (UYDU-) UNIV DUKE.
XX
PI  Haynes BF, Liao H;
XX
DR  WPI; 2001-488827/53.
XX
PT  Vaccine for immunizing against human immunodeficiency virus has mixture
PT  or linear array of peptides comprising immunodominant T-helper epitopes
PT  and major histocompatibility complex cytotoxic T-lymphocyte epitopes.
XX
PS  Disclosure; Page 27; 33pp; English.
XX
CC  The present invention relates to human immunodeficiency virus (HIV) and
CC  in particular to a human leukocyte antigen (HLA)-based HIV vaccine. The
CC  vaccine comprises a mixture or linear array of peptides, or its variants,
CC  where the peptides contain immunodominant T-helper (Th) epitopes and
CC  major histocompatibility complex (MHC) cytotoxic T- lymphocyte (CTL)
CC  epitopes and the linear array of peptides are preferably expressed in
CC  modified Vaccinia ankara. The vaccine is useful for immunising a patient
CC  against HIV and focuses immune response on many dominant and subdominant
CC  CTL epitopes of HIV. DNA or live vectors with linear arrays of CTL
CC  epitopes can be used as either primes or boosts of peptides or of each
CC  other to optimally give CTL anti-HIV responses. The vaccine induces
CC  salutary anti-HIV immune responses. AAU12447-AAU12558 represent the amino
CC  acid sequences of the Th-CTL epitopes and HIV immunogenic peptides used
CC  in the invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ  Sequence 15 AA;

Query Match      100.0%; Score 48; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1  LWNWFNI 7
      |||||
DB      3  LWNWFNI 9

RESULT 4
AAU12524
ID  AAU12524 standard; peptide; 15 AA.
XX
AC  AAU12524;
XX
KW  Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW  human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW  major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW  Vaccinia ankara.
XX
DT  11-SEP-2003 (revised)
DT  27-SEP-2001 (first entry)
XX
XX

```


DE Human HIV-1 Th-CTL overlapping epitope #31.
 XX Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
 KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
 KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
 KW Vaccinia ankara.
 XX Homo sapiens.
 OS Human immunodeficiency virus 1.
 XX WO200156355-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US003540.
 XX 04-FEB-2000; 2000US-00497497.
 XX (UYDU-) UNIV DUKE.
 XX Haynes BF, Liao H;
 XX WPI; 2001-488827/53.
 XX Vaccine for immunizing against human immunodeficiency virus has mixture
 PT or linear array of peptides comprising immunodominant T-helper epitopes
 PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes.
 XX Claim 1; Page 26; 33pp; English.
 XX The present invention relates to human immunodeficiency virus (HIV) and
 CC in particular to a human leukocyte antigen (HLA)-based HIV vaccine. The
 CC vaccine comprises a mixture or linear array of peptides, or its variants,
 CC where the peptides contain immunodominant T-helper (Th) epitopes and
 CC major histocompatibility complex (MHC) cytotoxic T-lymphocyte (CTL)
 CC epitopes and the linear array of peptides are preferably expressed in
 CC modified vaccinia ankara. The vaccine is useful for immunising a patient
 CC against HIV and focuses immune response on many dominant and subdominant
 CC CTL epitopes of HIV. DNA or live vectors with linear arrays of CTL
 CC epitopes can be used as either primes or boosts of peptides or of each
 CC other to optimally give CTL anti-HIV responses. The vaccine induces
 CC salutary anti-HIV immune responses. AAU12447-AAU12558 represent the amino
 CC acid sequences of the Th-CTL epitopes and HIV immunogenic peptides used
 CC in the invention. (Updated on 11-SEP-2003 to standardise OS field)
 XX Sequence 15 AA;
 SQ

Query Match 100.0%; Score 48; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.66;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFNI 7
 DB 3 LWNWFNI 9
 |||||

RESULT 5
 AAW07920
 ID AAW07920 standard; peptide; 20 AA.
 XX
 AC AAW07920;
 XX
 DT 16-OCT-2003 (revised)
 DT 31-JAN-1997 (first entry)
 XX
 DE gp41 peptide 53.
 XX HIV; gp120; HIV-IIIB strain; HIV-1 transmission; foetal transmission;
 KW neutralising antibody; passive immunisation; anti-idiotypic antibody;
 KW gp41; vaccine; active immunotherapy.
 XX Human immunodeficiency virus 1.
 OS

PN US5556744-A.
 XX 17-SEP-1996.
 XX 24-MAR-1994; 94US-00218025.
 XX 29-MAY-1992; 92US-00891451.
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 XX Williams WV, Weiner DB, Ugen KB;
 PI WPI; 1996-432980/43.
 XX Determining the likelihood of maternal transmission of HIV-1 to foetus -
 PT by measuring maternal reactivity with specific gp120 and gp41 derived
 PT peptide(s), also used for diagnosing HIV in infants.
 XX Claim 1; Col 115; 63pp; English.
 XX AAW07918-W07928 represent HIV gp41 peptides that can be used in the
 CC method of the invention. These sequences are derived from the HIV-IIIB
 CC strain deposited as ATCC CRL 8543. The method of the invention is for
 CC determining whether or not a mother will transmit HIV-1 to a foetus. The
 CC method comprises incubating a sample from the HIV-infected mother, with a
 CC collection of HIV peptides. The HIV peptides includes at least one of
 CC these sequences, and at least one HIV gp120 derived peptide (see AAW07909
 CC -W07917). The number of peptides that react with the sample is
 CC determined, and this number is compared with a standard that shows
 CC pattern reactivity for a patient of transmission status. A non-
 CC transmissible HIV sample is indicated if the test sample reacts with twice
 CC as many peptides as the standard. The method detects the presence of
 CC neutralising antibodies that protect against mother to infant
 CC transmission of HIV. These sequences can also be used in vaccines to
 CC protect against transmission. Antibodies against these sequences can be
 CC used for passive immunisation, and to generate anti-idiotypic antibodies
 CC for use in vaccines or active immunotherapy. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX Sequence 20 AA;
 SQ

Query Match 100.0%; Score 48; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.89;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFNI 7
 DB 8 LWNWFNI 14
 |||||

RESULT 6
 AAP82469
 ID AAP82469 standard; protein; 23 AA.
 XX
 AC AAP82469;
 XX
 DT 25-MAR-2003 (revised)
 DT 12-NOV-1990 (first entry)
 XX
 DE Peptide component of AIDS vaccine.
 XX AIDS vaccine; T-cells.
 XX Synthetic.
 XX EP273716-A.
 XX 06-JUL-1988.
 XX 23-DEC-1987; 87EP-00311391.
 XX 30-DEC-1986; 86US-00947935.

PR 12-FEB-1987; 87US-00014430.
 XX (USDC) US SEC OF COMMERCE.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX Delisi C, Margalit H, Cornette JL, Ouyang CS;
 XX WPI; 1988-184640/27.
 XX Synthetic peptide(s) as vaccines for AIDS - selected from peptide regions
 PT which can fold as a maximally amphipathic helix recognised by I cells.
 XX Claim 9; Page 10; 16pp; English.
 XX This peptide is a component of an AIDS vaccine. It can fold as a
 CC maximally amphipathic helix and is recognised by T-cells immune to the
 CC AIDS virus envelope protein. See also AAP82462-68 and AAP82470-79.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 23 AA;
 SQ
 Query Match 100.0%; Score 48; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWNWFNI 7
 DB 11 LWNWFNI 17
 RESULT 7
 ID AAW07974 standard; peptide; 23 AA.
 XX
 AC AAW07974;
 XX
 DT 16-OCT-2003 (revised)
 DT 03-FEB-1997 (first entry)
 XX
 DE HIV peptide 41-21.
 XX
 KW HIV; gp120; HIV-IIIB strain; HIV-1 transmission; foetal transmission;
 KW neutralising antibody; passive immunisation; anti-idiotypic antibody;
 KW gp41; vaccine; active immunotherapy.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5556744-A.
 XX
 PD 17-SEP-1996.
 XX
 PF 24-MAR-1994; 94US-00218025.
 XX
 PR 29-MAY-1992; 92US-00891451.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 XX
 PI Williams WV, Weiner DB, Ugen KE;
 XX
 DR WPI; 1996-432980/43.
 XX
 XX Determining the likelihood of maternal transmission of HIV-1 to foetus -
 PT by measuring maternal reactivity with specific gp120 and gp41 derived
 PT peptide(s), also used for diagnosing HIV in infants.
 XX
 XX Example 6; Col 95-96; 63pp; English.
 PS
 CC AAW07956-W07979 represent HIV peptides that can be used in the method of
 CC the invention. The method of the invention is for determining whether or
 CC not a mother will transmit HIV-1 to a foetus. The method comprises
 CC incubating a sample from the HIV-infected mother, with a collection of
 CC HIV peptides. The HIV peptides includes at least one HIV gp120 derived

CC peptide (see AAW07909-W07917), and at least one HIV gp41 derived peptide
 CC (see AAW07918-W07928). The number of peptides that react with the sample
 CC is determined, and this number is compared with a standard that shows
 CC pattern reactivity for a patient of transmission status. A non-
 CC transmissible HIV sample is indicated if the test sample reacts with twice
 CC as many peptides as the standard. The method detects the presence of
 CC neutralising antibodies that protect against mother to infant
 CC transmission of HIV. These sequences can also be used in vaccines to
 CC protect against transmission. Antibodies against these sequences can be
 CC used for passive immunisation, and to generate anti-idiotypic antibodies
 CC for use in vaccines or active immunotherapy. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX Sequence 23 AA;
 SQ
 Query Match 100.0%; Score 48; DB 2; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWNWFNI 7
 DB 9 LWNWFNI 15
 RESULT 8
 ID AAW12483 standard; peptide; 23 AA.
 XX
 AC AAW12483;
 XX
 DT 27-SEP-2001 (first entry)
 XX
 DE HIV Th-CTL peptide prototype vaccine immunogen #20.
 XX
 KW Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
 KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
 KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
 KW Vaccinia ankara.
 XX
 OS Homo sapiens.
 XX
 PN WO200156355-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US003540.
 XX
 PR 04-FEB-2000; 2000US-00497497.
 XX
 PA (UYDU-) UNIV DUKE.
 XX
 PI Haynes BF, Liao H;
 XX
 DR WPI; 2001-488827/53.
 XX
 PT Vaccine for immunizing against human immunodeficiency virus has mixture
 PT or linear array of peptides comprising immunodominant T-helper epitopes
 PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes.
 XX
 PS Claim 1; Page 24; 33pp; English.
 XX
 CC The present invention relates to human immunodeficiency virus (HIV) and
 CC in particular to a human leukocyte antigen (HLA)-based HIV vaccine. The
 CC vaccine comprises a mixture or linear array of peptides, or its variants,
 CC where the peptides contain immunodominant T-helper (Th) epitopes and
 CC major histocompatibility complex (MHC) cytotoxic T- lymphocyte (CTL)
 CC epitopes and the linear array of peptides are preferably expressed in
 CC modified Vaccinia ankara. The vaccine is useful for immunising a patient
 CC against HIV and focuses immune response on many dominant and subdominant
 CC CTL epitopes of HIV. DNA or live vectors with linear arrays of CTL
 CC epitopes can be used as either primes or boosts of peptides or of each
 CC other to optimally give CTL anti-HIV responses. The vaccine induces
 CC salutary anti-HIV immune responses. AAW12447-AAU12558 represent the amino

CC acid sequences of the Th-CTL epitopes and HIV immunogenic peptides used
 CC in the invention
 XX
 SQ Sequence 23 AA;

Query Match 100.0%; Score 48; DB 4; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
 |||||
 Db 3 LMNWFNI 9

RESULT 9

AAU12488
 ID AAU12488 standard; peptide; 24 AA.

AC AAU12488;

XX 27-SEP-2001 (first entry)

DE HIV Th-CTL peptide prototype vaccine immunogen #25.

XX Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
 KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
 KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
 KW Vaccinia ankara.

XX Homo sapiens.

OS WO200156355-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US003540.

XX 04-FEB-2000; 2000US-00497497.

XX (UYDU-) UNIV DUKE.

XX Haynes BF, Liao H;

XX WPI; 2001-488827/53.

XX Vaccine for immunizing against human immunodeficiency virus has mixture
 PT or linear array of peptides comprising immunodominant T-helper epitopes
 PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes.

XX Claim 1; Page 25; 33pp; English.

XX The present invention relates to human immunodeficiency virus (HIV) and
 CC in particular to a human leukocyte antigen (HLA)-based HIV vaccine. The
 CC vaccine comprises a mixture or linear array of peptides, or its variants,
 CC where the peptides contain immunodominant T-helper (Th) epitopes and
 CC major histocompatibility complex (MHC) cytotoxic T-lymphocyte (CTL)
 CC epitopes and the linear array of peptides are preferably expressed in
 CC modified Vaccinia ankara. The vaccine is useful for immunising a patient
 CC against HIV and focuses immune response on many dominant and subdominant
 CC CTL epitopes of HIV. DNA or live vectors with linear arrays of CTL
 CC epitopes can be used as either primes or boosts of peptides or of each
 CC other to optimally give CTL anti-HIV responses. The vaccine induces
 CC salutary anti-HIV immune responses. AAU12447-AAU12558 represent the amino
 CC acid sequences of the Th-CTL epitopes and HIV immunogenic peptides used
 CC in the invention

XX SQ Sequence 24 AA;

Query Match 100.0%; Score 48; DB 4; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7

Db 3 LMNWFNI 9

RESULT 10

AAU70251
 ID AAU70251 standard; peptide; 24 AA.

XX AAU70251;

XX 14-FEB-2002 (first entry)

XX HIV TH-CTL P17 epitopes (A2 variants) prototype vaccine D-TH/L-CTL.

XX HIV; human immunodeficiency virus; AIDS; SIV; anti-HIV; virucide;
 KW acquired immunodeficiency syndrome; simian immunodeficiency virus;
 KW vaccine; cytotoxic T cell response inducer; Th; human leucocyte antigen;
 KW immunodominant T-helper epitope; cytotoxic T lymphocyte; CTL; HLA.

XX Human immunodeficiency virus 1.
 OS Synthetic.

XX US2001036461-A1.

XX 01-NOV-2001.

XX 05-FEB-2001; 2001US-00775805.

XX 04-FEB-2000; 2000US-00497497.

XX (HAYN/) HAYNES B F.

XX (LIAO/) LIAO H.

XX Haynes BF, Liao H;

XX WPI; 2002-010132/01.

XX Vaccine containing overlapping Th-cytotoxic T lymphocyte epitopes, useful
 PT for protecting against human immune deficiency virus.

XX Example 1; Page 6; lipp; English.

XX The invention relates to a vaccine comprising a mixture, or linear array,
 CC of specific peptides that are human immunodeficiency virus (HIV-1) Th
 CC (immunodominant T-helper epitope)-cytotoxic T lymphocyte (CTL), or their
 CC variants. Peptides are also described which contain SIV (simian
 CC immunodeficiency virus) derived epitopes. The vaccine is used to protect
 CC against infection by HIV. The vaccines are based on an analysis of HLA
 CC (human leucocyte antigen) alleles present in subject populations targeted
 CC for vaccination and the commonest variants of HIV in the locality. They
 CC contain sufficient immunogenic epitopes to ensure effective presentation
 CC by almost all members of the target population. The present sequence is
 CC an HIV-1 epitope based prototype vaccine of the invention

XX SQ Sequence 24 AA;

Query Match 100.0%; Score 48; DB 5; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7

Db 3 LMNWFNI 9

RESULT 11

AAU12479
 ID AAU12479 standard; peptide; 32 AA.

XX AAU12479;

XX 27-SEP-2001 (first entry)

XX

XX Peptide-simulator of human immunodeficiency type-1 virus protein gp41
 PT conservative epitope recognized by virus-neutralizing monoclonal antibody
 PT 2f5 (variants).

XX Disclosure; Fig 3; 6pp; Russian.

XX The present invention describes 12 peptide-simulators (ABB09187 to
 CC ABB09198) of HIV-1 (human immunodeficiency virus type 1) protein gp41
 CC conservative epitope prepared by affinity selection from phage peptide
 CC libraries. These peptides differ from the HIV-1 protein gp41 conservative
 CC epitope sequence but retain the ability to bind with the virus-
 CC neutralising monoclonal antibodies 2F5. The invention also describes a
 CC vaccine against HIV based on the peptides. The peptides can be used in
 CC biotechnology, virology, and vaccines. The present sequence represents an
 CC HIV-1 gp160 amino acid sequence given in the exemplification of the
 CC present invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 33 AA;

Query Match 100.0%; Score 48; DB 5; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFEI 7
 Db 18 LWNWFEI 24
 |||||

RESULT 14

AAW17022
 ID AAW17022 standard; peptide; 36 AA.

XX AAW17022;

DT 01-JUL-1997 (first entry)

DE DP-178-derived peptide, T636, having high antiviral activity.

XX HIV; SIV; human; simian immunodeficiency virus; glycoprotein 41;
 KW transmembrane; antiviral; DP-178; DP-107; vaccine; inhibition;
 KW replication; transmission.

XX Synthetic.

XX WO9640191-A1.

PD 19-DEC-1996..

XX 06-JUN-1996; 96WO-US009499.

XX 07-JUN-1995; 95US-00481957.

XX (TRIM-) TRIMERIS INC.

XX Johnson RM, Lambert DM;

XX WPI; 1997-099886/09.

XX Compens. contg. DP-178 or DP-107 in combination with other therapeutic
 PT agent - useful for treatment of HIV infection, esp. by inhibiting
 PT replication or transmission of HIV.

XX Claim 8; Fig 5A (III-IV); 84pp; English.

XX AAW17021-W17028 are DP-178-derived peptides that have high antiviral
 CC activity. DP-178 is a peptide corresponding to residues 638-673 of HIV
 CC type 1 glycoprotein 41 (gp41) transmembrane protein. DP-178 and its
 CC derivatives/homologues are used in combination with a therapeutic agent,
 CC e.g. a reverse transcriptase, viral protease, cytokine, glycosylation or
 CC viral mRNA processing inhibitor or a nucleoside inhibitor. The peptides
 CC work by inhibiting viral replication or inhibiting transmission. They may
 CC also be used in vaccines for protecting against HIV infection

XX SQ Sequence 36 AA;

Query Match 100.0%; Score 48; DB 2; Length 36;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFEI 7
 Db 28 LWNWFEI 34
 |||||

RESULT 15

AAW89191
 ID AAW89191 standard; peptide; 36 AA.

XX AAW89191;

XX 23-MAY-2000 (first entry)

XX Core polypeptide fragment T No. 629.

XX Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
 KW HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;
 KW anti-fusogenic; differentiation factor; interleukin; interferon;
 KW colony stimulating factor; hormone; angiogenic factor.

XX Unidentified.

XX WO9959615-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-US011219.

XX 20-MAY-1998; 98US-00082279.

XX (TRIM-) TRIMERIS INC.

XX Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;

XX WPI; 2000-136792/12.

XX A new hybrid polypeptide with enhanced pharmacokinetic properties
 PT comprises enhancer sequence.

XX Disclosure; Page 31; 124pp; English.

XX The invention relates to hybrid polypeptides comprising enhancer peptide
 CC sequence linked to core polypeptides. The enhancer polypeptides are
 CC derived from various retroviral envelope (gp41) protein sequences,
 CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
 CC pharmacokinetic properties such as increasing the half-life of any core
 CC polypeptide that they are linked to. The core polypeptides are any
 CC polypeptide that may be introduced into a living system and that can
 CC function as a pharmacologically useful peptide for the treatment or
 CC prevention of a disease. The core polypeptides are bioactive peptides
 CC selected from a growth factor, cytokine, differentiation factor,
 CC interleukin, interferon, colony stimulating factor, hormone or angiogenic
 CC factor. The peptides of the invention can be used for inhibiting viral
 CC infection and can be used in anti-viral and anti-fusogenic treatments.
 CC Sequences AAW8651-Y90055 represent core polypeptide fragments that can
 CC be used in the invention. Some sequences among those indicated also
 CC comprise enhancer fragments at terminal ends and form hybrid polypeptides

XX SQ Sequence 36 AA;

Query Match 100.0%; Score 48; DB 3; Length 36;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFEI 7
 |||||

Db 21 LMNWFNI 27

Search completed: February 25, 2004, 15:41:08
Job time : 12.3228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:37:04 ; Search time 2.79114 Seconds
(without alignments)
129.475 Million cell updates/sec

Title: US-09-877-606-7
Perfect score: 48
Sequence: 1 LMWFWNI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 48 | 100.0 | 20 | 1 | US-08-218-025A-7 |
| 2 | 48 | 100.0 | 23 | 1 | US-08-218-025A-152 |
| 3 | 48 | 100.0 | 36 | 3 | US-08-484-223B-232 |
| 4 | 48 | 100.0 | 36 | 3 | US-08-484-223B-233 |
| 5 | 48 | 100.0 | 36 | 3 | US-09-082-279B-549 |
| 6 | 48 | 100.0 | 36 | 3 | US-09-082-279B-550 |
| 7 | 48 | 100.0 | 36 | 3 | US-09-082-279B-551 |
| 8 | 48 | 100.0 | 36 | 3 | US-09-082-279B-552 |
| 9 | 48 | 100.0 | 36 | 3 | US-09-082-279B-553 |
| 10 | 48 | 100.0 | 36 | 3 | US-09-082-279B-554 |
| 11 | 48 | 100.0 | 36 | 3 | US-09-082-279B-555 |
| 12 | 48 | 100.0 | 36 | 3 | US-09-082-279B-556 |
| 13 | 48 | 100.0 | 36 | 3 | US-09-082-279B-557 |
| 14 | 48 | 100.0 | 36 | 3 | US-09-082-279B-558 |
| 15 | 48 | 100.0 | 36 | 3 | US-09-082-279B-559 |
| 16 | 48 | 100.0 | 36 | 3 | US-09-082-279B-560 |
| 17 | 48 | 100.0 | 36 | 4 | US-08-474-349A-401 |
| 18 | 48 | 100.0 | 36 | 4 | US-08-474-349A-402 |
| 19 | 48 | 100.0 | 36 | 4 | US-08-474-349A-403 |
| 20 | 48 | 100.0 | 36 | 4 | US-08-474-349A-404 |
| 21 | 48 | 100.0 | 36 | 4 | US-08-474-349A-405 |
| 22 | 48 | 100.0 | 36 | 4 | US-08-474-349A-406 |
| 23 | 48 | 100.0 | 36 | 4 | US-08-474-349A-407 |
| 24 | 48 | 100.0 | 36 | 4 | US-08-474-349A-408 |
| 25 | 48 | 100.0 | 36 | 4 | US-08-474-349A-409 |
| 26 | 48 | 100.0 | 36 | 4 | US-08-474-349A-410 |
| 27 | 48 | 100.0 | 36 | 4 | US-08-474-349A-411 |

| | | | | | | |
|----|----|-------|----|---|--------------------|-------------------|
| 28 | 48 | 100.0 | 36 | 4 | US-08-474-349A-412 | Sequence 412, App |
| 29 | 48 | 100.0 | 36 | 4 | US-09-315-304B-549 | Sequence 549, App |
| 30 | 48 | 100.0 | 36 | 4 | US-09-315-304B-550 | Sequence 550, App |
| 31 | 48 | 100.0 | 36 | 4 | US-09-315-304B-551 | Sequence 551, App |
| 32 | 48 | 100.0 | 36 | 4 | US-09-315-304B-552 | Sequence 552, App |
| 33 | 48 | 100.0 | 36 | 4 | US-09-315-304B-553 | Sequence 553, App |
| 34 | 48 | 100.0 | 36 | 4 | US-09-315-304B-554 | Sequence 554, App |
| 35 | 48 | 100.0 | 36 | 4 | US-09-315-304B-555 | Sequence 555, App |
| 36 | 48 | 100.0 | 36 | 4 | US-09-315-304B-556 | Sequence 556, App |
| 37 | 48 | 100.0 | 36 | 4 | US-09-315-304B-557 | Sequence 557, App |
| 38 | 48 | 100.0 | 36 | 4 | US-09-315-304B-558 | Sequence 558, App |
| 39 | 48 | 100.0 | 36 | 4 | US-09-315-304B-559 | Sequence 559, App |
| 40 | 48 | 100.0 | 36 | 4 | US-09-315-304B-560 | Sequence 560, App |
| 41 | 48 | 100.0 | 36 | 4 | US-08-973-952-56 | Sequence 56, Appl |
| 42 | 48 | 100.0 | 36 | 4 | US-08-973-952-57 | Sequence 57, Appl |
| 43 | 48 | 100.0 | 36 | 4 | US-08-973-952-58 | Sequence 58, Appl |
| 44 | 48 | 100.0 | 36 | 4 | US-09-834-784-549 | Sequence 549, App |
| 45 | 48 | 100.0 | 36 | 4 | US-09-834-784-550 | Sequence 550, App |

ALIGNMENTS

RESULT 1
US-08-218-025A-7
; Sequence 7, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Ugen, Kenneth E.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; TITLE OF INVENTION: Methods and Compositions for Treating HIV Infected Patients
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 5556744ristown Road
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,025A
FILING DATE: 24-MAR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,451
FILING DATE: 29-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST33A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-218-025A-7

Query Match 100.0%; Score 48; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0

QY 1 LMWFWNI 7

| | | | | |
|--|---|---|--|--|
| | / | N U M B E R O F S E Q U E N C E S : | 245 | |
| | / | C O R R E S P O N D E N C E A D D R E S S : | | |
| | / | A D D R E S S E E: | Pennie & Edmonds LLP | |
| | / | S T R E E T: | 1155 Avenue of the Americas | |
| | / | C I T Y: | New York | |
| | / | S T A T E: | New York | |
| | / | C O U N T R Y: | USA | |
| | / | Z I P: | 10036-2711 | |
| | / | C O M P U T E R R E A D A B L E F O R M: | | |
| | / | M E D I U M T Y P E: | Floppy disk | |
| | / | C O M P U T E R: | IBM PC compatible | |
| | / | O P E R A T I N G S Y S T E M: | PC-DOS/MS-DOS | |
| | / | S O F T W A R E: | PatentIn Release #1.0, Version #1.30 | |
| | / | T I T L E O F I N V E N T I O N: | and Treating Certain HIV Infected Patients | |
| | / | N U M B E R O F S E Q U E N C E S: | 197 | |
| | / | C O R R E S P O N D E N C E A D D R E S S: | | |
| | / | A D D R E S S E E: | Howsen and Howson | |
| | / | S T R E E T: | P.O. Box 457, 321 No. 5556744ristown Road | |
| | / | C I T Y: | Spring House | |
| | / | S T A T E: | Pennsylvania | |
| | / | C O U N T R Y: | U.S.A. | |
| | / | Z I P: | 19477 | |
| | / | C O M P U T E R R E A D A B L E F O R M: | | |
| | / | M E D I U M T Y P E: | Floppy disk | |
| | / | C O M P U T E R: | IBM PC compatible | |
| | / | O P E R A T I N G S Y S T E M: | PC-DOS/MS-DOS | |
| | / | S O F T W A R E: | PatentIn Release #1.0, Version #1.25 | |
| | / | C U R R E N T A P P L I C A T I O N N U M B E R: | US/08/218,025A | |
| | / | F I L I N G D A T E: | 24-MAR-1994 | |
| | / | C L A S S I F I C A T I O N: | 424 | |
| | / | P R I O R A P P L I C A T I O N I N F O R M A T I O N: | | |
| | / | A P P L I C A T I O N N U M B E R: | US 07/891,451 | |
| | / | F I L I N G D A T E: | 29-MAY-1992 | |
| | / | A T T O R N E Y/ A G E N T I N F O R M A T I O N: | | |
| | / | N A M E: | Bak, Mary E. | |
| | / | R E G I S T R A T I O N N U M B E R: | 31,215 | |
| | / | R E F E R E N C E/ D O C K E T N U M B E R: | WST33A | |
| | / | T E L E C O M M U N I C A T I O N I N F O R M A T I O N: | | |
| | / | T E L E P H O N E: | (215) 540-9206 | |
| | / | T E L E F A X: | (215) 540-5818 | |
| | / | I N F O R M A T I O N F O R S E Q I D N O: | 152 | |
| | / | S E Q U E N C E C H A R A C T E R I S T I C S: | | |
| | / | L E N G T H: | 23 amino acids | |
| | / | T Y P E: | amino acid | |
| | / | T O P O L O G Y: | unknown | |
| | / | M O L E C U L E T Y P E: | peptide | |
| | / | U S - 0 8 - 2 1 8 - 0 2 5 A - 1 5 2 | | |
| | / | Q u e r y M a t c h | 100.0%; Score 48; DB 1; Length 23; | |
| | / | B e s t L o c a l S i m i l a r i t y | 100.0%; Pred.No. 0.19; | |
| | / | M a t c h e s | 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| | / | D b | 1 LWNWFNI 7 | |
| | / | | | |
| | / | | 30 LWNWFNI 36 | |
| | / | R E S U L T | 4 | |
| | / | U S - 0 8 - 4 8 4 - 2 2 3 B - 2 3 3 | | |
| | / | S e q u e n c e | 233 Application US/08484223B | |
| | / | P a t e n t | No. 6020459 | |
| | / | G E N E R A L I N F O R M A T I O N: | | |
| | / | A P P L I C A N T: | Bolognesi, Dani P. | |
| | / | A P P L I C A N T: | Matthews, Thomas J. | |
| | / | A P P L I C A N T: | Wild, Carl T. | |
| | / | A P P L I C A N T: | Barney, Shawn O. | |
| | / | A P P | | |

[illegible][illegible]

;
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 869-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 233:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-484-223B-233

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFNI 7
Db 29 LNNWFNI 35

RESULT 5
US-09-082-279B-549
; Sequence 549, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 549
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-549

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFNI 7
Db 19 LNNWFNI 25

RESULT 6
US-09-082-279B-550
; Sequence 550, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED

;
; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 550
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-550

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFNI 7
Db 20 LNNWFNI 26

RESULT 7
US-09-082-279B-551
; Sequence 551, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 551
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-551

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFNI 7
Db 21 LNNWFNI 27

RESULT 8
US-09-082-279B-552
; Sequence 552, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515

```
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 552
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Core polypeptide
US-09-082-2798-552

Query Match      100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LWNWFNI 7
Db      22 LWNWFNI 28

RESULT 9
US-09-082-2798-553
/ Sequence 553, Application US/090822798
/ Patent No. 6258782
/ GENERAL INFORMATION:
/ APPLICANT: Barney, Shawn
/ APPLICANT: Guthrie, Kelly
/ APPLICANT: Merutka, Gene
/ APPLICANT: Anwer, Mohamed
/ APPLICANT: Lambert, Dennis
/ TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
/ FILE REFERENCE: 7872-043
/ CURRENT APPLICATION NUMBER: US/09/082,2798
/ CURRENT FILING DATE: 1998-05-20
/ NUMBER OF SEQ ID NOS: 1515
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 553
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Core polypeptide
US-09-082-2798-553

Query Match      100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LWNWFNI 7
Db      23 LWNWFNI 29

RESULT 10
US-09-082-2798-554
/ Sequence 554, Application US/090822798
/ Patent No. 6258782
/ GENERAL INFORMATION:
/ APPLICANT: Barney, Shawn
/ APPLICANT: Guthrie, Kelly
/ APPLICANT: Merutka, Gene
/ APPLICANT: Anwer, Mohamed
/ APPLICANT: Lambert, Dennis
/ TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
/ FILE REFERENCE: 7872-043
/ CURRENT APPLICATION NUMBER: US/09/082,2798
/ CURRENT FILING DATE: 1998-05-20
/ NUMBER OF SEQ ID NOS: 1515
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 554
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
US-09-082-2798-554

Query Match      100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LWNWFNI 7
Db      25 LWNWFNI 31

RESULT 11
US-09-082-2798-555
/ Sequence 555, Application US/090822798
/ Patent No. 6258782
/ GENERAL INFORMATION:
/ APPLICANT: Barney, Shawn
/ APPLICANT: Guthrie, Kelly
/ APPLICANT: Merutka, Gene
/ APPLICANT: Anwer, Mohamed
/ APPLICANT: Lambert, Dennis
/ TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
/ FILE REFERENCE: 7872-043
/ CURRENT APPLICATION NUMBER: US/09/082,2798
/ CURRENT FILING DATE: 1998-05-20
/ NUMBER OF SEQ ID NOS: 1515
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 555
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Core polypeptide
US-09-082-2798-555

Query Match      100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LWNWFNI 7
Db      24 LWNWFNI 30

RESULT 12
US-09-082-2798-556
/ Sequence 556, Application US/090822798
/ Patent No. 6258782
/ GENERAL INFORMATION:
/ APPLICANT: Barney, Shawn
/ APPLICANT: Guthrie, Kelly
/ APPLICANT: Merutka, Gene
/ APPLICANT: Anwer, Mohamed
/ APPLICANT: Lambert, Dennis
/ TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
/ FILE REFERENCE: 7872-043
/ CURRENT APPLICATION NUMBER: US/09/082,2798
/ CURRENT FILING DATE: 1998-05-20
/ NUMBER OF SEQ ID NOS: 1515
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 556
/ LENGTH: 36
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Core polypeptide
US-09-082-2798-556

Query Match      100.0%; Score 48; DB 3; Length 36;
```

```
Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 LMNWFNI 7
Db 26 LMNWFNI 32

RESULT 13
US-09-082-279B-557
; Sequence 557, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 557
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-557

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 27 LMNWFNI 33

RESULT 14
US-09-082-279B-558
; Sequence 558, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 558
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-558

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 27 LMNWFNI 33

RESULT 15
US-09-082-279B-559
; Sequence 559, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 559
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-559

Query Match 100.0%; Score 48; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 29 LMNWFNI 35

Search completed: February 25, 2004, 15:46:55
Job time : 2.79114 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:44:50 ; Search time 5.8481 Seconds
(without alignments)
252.744 Million cell updates/sec

Title: US-09-877-606-7

Perfect score: 48
Sequence: 1 LMNWFNI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21113259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 48 | 100.0 | 36 | 14 | US-10-252-136-56 |
| 2 | 48 | 100.0 | 36 | 14 | US-10-252-136-57 |
| 3 | 48 | 100.0 | 36 | 14 | US-10-252-136-58 |
| 4 | 48 | 100.0 | 36 | 14 | US-10-351-641-549 |
| 5 | 48 | 100.0 | 36 | 14 | US-10-351-641-550 |
| 6 | 48 | 100.0 | 36 | 14 | US-10-351-641-551 |
| 7 | 48 | 100.0 | 36 | 14 | US-10-351-641-552 |
| 8 | 48 | 100.0 | 36 | 14 | US-10-351-641-553 |
| 9 | 48 | 100.0 | 36 | 14 | US-10-351-641-554 |
| 10 | 48 | 100.0 | 36 | 14 | US-10-351-641-555 |
| 11 | 48 | 100.0 | 36 | 14 | US-10-351-641-556 |
| 12 | 48 | 100.0 | 36 | 14 | US-10-351-641-557 |
| 13 | 48 | 100.0 | 36 | 14 | US-10-351-641-558 |
| 14 | 48 | 100.0 | 36 | 14 | US-10-351-641-559 |
| 15 | 48 | 100.0 | 36 | 14 | US-10-351-641-560 |

| | | | | | | |
|----|----|-------|-----|----|-------------------|--------------------|
| 16 | 48 | 100.0 | 38 | 14 | US-10-351-641-776 | Sequence 776, Appl |
| 17 | 48 | 100.0 | 39 | 14 | US-10-351-641-777 | Sequence 777, Appl |
| 18 | 48 | 100.0 | 40 | 14 | US-10-351-641-778 | Sequence 778, Appl |
| 19 | 48 | 100.0 | 40 | 14 | US-10-351-641-781 | Sequence 781, Appl |
| 20 | 48 | 100.0 | 49 | 14 | US-10-252-136-13 | Sequence 13, Appl |
| 21 | 48 | 100.0 | 56 | 9 | US-09-779-451-4 | Sequence 4, Appl |
| 22 | 48 | 100.0 | 177 | 14 | US-10-040-349B-2 | Sequence 2, Appl |
| 23 | 48 | 100.0 | 200 | 14 | US-10-263-103-25 | Sequence 25, Appl |
| 24 | 48 | 100.0 | 200 | 15 | US-10-438-691-8 | Sequence 8, Appl |
| 25 | 48 | 100.0 | 221 | 14 | US-10-059-271-84 | Sequence 84, Appl |
| 26 | 48 | 100.0 | 232 | 14 | US-10-059-271-81 | Sequence 81, Appl |
| 27 | 48 | 100.0 | 254 | 14 | US-10-059-271-82 | Sequence 82, Appl |
| 28 | 48 | 100.0 | 256 | 14 | US-10-059-271-97 | Sequence 97, Appl |
| 29 | 48 | 100.0 | 268 | 9 | US-09-854-816-9 | Sequence 9, Appl |
| 30 | 48 | 100.0 | 268 | 9 | US-09-854-816-10 | Sequence 10, Appl |
| 31 | 48 | 100.0 | 268 | 9 | US-09-854-816-16 | Sequence 16, Appl |
| 32 | 48 | 100.0 | 268 | 9 | US-09-854-816-17 | Sequence 17, Appl |
| 33 | 48 | 100.0 | 268 | 9 | US-09-854-816-18 | Sequence 18, Appl |
| 34 | 48 | 100.0 | 268 | 9 | US-09-854-816-19 | Sequence 19, Appl |
| 35 | 48 | 100.0 | 268 | 9 | US-09-854-816-41 | Sequence 41, Appl |
| 36 | 48 | 100.0 | 269 | 9 | US-09-854-816-12 | Sequence 12, Appl |
| 37 | 48 | 100.0 | 269 | 9 | US-09-854-816-31 | Sequence 31, Appl |
| 38 | 48 | 100.0 | 269 | 9 | US-09-854-816-43 | Sequence 43, Appl |
| 39 | 48 | 100.0 | 269 | 9 | US-09-854-816-44 | Sequence 44, Appl |
| 40 | 48 | 100.0 | 269 | 9 | US-09-854-816-46 | Sequence 46, Appl |
| 41 | 48 | 100.0 | 344 | 14 | US-10-040-349B-1 | Sequence 1, Appl |
| 42 | 48 | 100.0 | 345 | 9 | US-09-779-451-8 | Sequence 8, Appl |
| 43 | 48 | 100.0 | 345 | 14 | US-10-026-741-49 | Sequence 49, Appl |
| 44 | 48 | 100.0 | 359 | 14 | US-10-214-670-58 | Sequence 58, Appl |
| 45 | 48 | 100.0 | 359 | 14 | US-10-214-670-62 | Sequence 62, Appl |

ALIGNMENTS

RESULT 1
US-10-252-136-56
; Sequence 56, Application US/10252136
; Publication No. US20030103998A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, M. Ross
; APPLICANT: Lambert, Dennis M.
; TITLE OF INVENTION: TREATMENT OF HUMAN IMMUNODEFICIENCY VIRUS AND OTHER
; TITLE OF INVENTION: VIRAL INFECTIONS
; TITLE OF INVENTION: USING COMBINATORY THERAPY
; FILE REFERENCE: 7872-036
; CURRENT APPLICATION NUMBER: US/10/252,136
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/08/973,952
; PRIOR FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 56
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-252-136-56

Query Match 100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMNWFNI 7
Db 27 LMNWFNI 33

RESULT 2
US-10-252-136-57
; Sequence 57, Application US/10252136
; Publication No. US20030103998A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, M. Ross

; APPLICANT: Lambert, Dennis M.
 ; TITLE OF INVENTION: TREATMENT OF HUMAN IMMUNODEFICIENCY VIRUS AND OTHER
 ; TITLE OF INVENTION: VIRAL INFECTIONS
 ; TITLE OF INVENTION: USING COMBINATORY THERAPY
 ; FILE REFERENCE: 7872-036
 ; CURRENT APPLICATION NUMBER: US/10/252,136
 ; CURRENT FILING DATE: 2002-09-20
 ; PRIOR APPLICATION NUMBER: US/08/973,952
 ; PRIOR FILING DATE: 1998-05-29
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 57
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus
 US-10-252-136-57

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFNI 7
 Db 29 LWNWFNI 35

RESULT 3
 US-10-252-136-58
 ; Sequence 58, Application US/10252136
 ; Publication No. US20030103998A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Johnson, M. Ross
 ; APPLICANT: Lambert, Dennis M.
 ; TITLE OF INVENTION: TREATMENT OF HUMAN IMMUNODEFICIENCY VIRUS AND OTHER
 ; TITLE OF INVENTION: VIRAL INFECTIONS
 ; TITLE OF INVENTION: USING COMBINATORY THERAPY
 ; FILE REFERENCE: 7872-036
 ; CURRENT APPLICATION NUMBER: US/10/252,136
 ; CURRENT FILING DATE: 2002-09-20
 ; PRIOR APPLICATION NUMBER: US/08/973,952
 ; PRIOR FILING DATE: 1998-05-29
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 58
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus
 US-10-252-136-58

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFNI 7
 Db 30 LWNWFNI 36

RESULT 4
 US-10-351-641-549
 ; Sequence 549, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.
 ; APPLICANT: Anwer, M.
 ; APPLICANT: Lambert, D.
 ; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
 ; TITLE OF INVENTION: PROPERTIES
 ; FILE REFERENCE: 7872-100
 ; CURRENT APPLICATION NUMBER: US/10/351,641
 ; CURRENT FILING DATE: 2003-01-24

; PRIOR APPLICATION NUMBER: 09/350,641
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/315,304
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/082,279
 ; PRIOR FILING DATE: 1998-05-20
 ; NUMBER OF SEQ ID NOS: 1757
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 549
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Core polypeptide
 US-10-351-641-549

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFNI 7
 Db 19 LWNWFNI 25

RESULT 5
 US-10-351-641-550
 ; Sequence 550, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.
 ; APPLICANT: Anwer, M.
 ; APPLICANT: Lambert, D.
 ; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
 ; TITLE OF INVENTION: PROPERTIES
 ; FILE REFERENCE: 7872-100
 ; CURRENT APPLICATION NUMBER: US/10/351,641
 ; CURRENT FILING DATE: 2003-01-24
 ; PRIOR APPLICATION NUMBER: 09/350,641
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/315,304
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/082,279
 ; PRIOR FILING DATE: 1998-05-20
 ; NUMBER OF SEQ ID NOS: 1757
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 550
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Core polypeptide
 US-10-351-641-550

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFNI 7
 Db 20 LWNWFNI 26

RESULT 6
 US-10-351-641-551
 ; Sequence 551, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.

APPLICANT: Anwer, M.
APPLICANT: Lambert, D.
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
FILE REFERENCE: 7872-100
CURRENT FILING DATE: 2003-01-24
CURRENT APPLICATION NUMBER: US/10/351,641
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/350,641
PRIOR FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/315,304
PRIOR FILING DATE: 1998-05-20
PRIOR APPLICATION NUMBER: 09/082,279
NUMBER OF SEQ ID NOS: 1757
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 551
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-10-351-641-551

Query Match 100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 LNNWFNI 7
Db 21 LNNWFNI 27

RESULT 7
US-10-351-641-552
Sequence 552, Application US/10351641
Publication No. US20030186874A1
GENERAL INFORMATION:
APPLICANT: Barney, S.
APPLICANT: Guthrie, K.
APPLICANT: Merutka, G.
APPLICANT: Anwer, M.
APPLICANT: Lambert, D.
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
FILE REFERENCE: 7872-100
CURRENT FILING DATE: 2003-01-24
CURRENT APPLICATION NUMBER: US/10/351,641
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/350,641
PRIOR FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/315,304
PRIOR FILING DATE: 1998-05-20
PRIOR APPLICATION NUMBER: 09/082,279
NUMBER OF SEQ ID NOS: 1757
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 552
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-10-351-641-552

Query Match 100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 LNNWFNI 7
Db 22 LNNWFNI 28

RESULT 8

Query Match 100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

US-10-351-641-553
Sequence 553, Application US/10351641
Publication No. US20030186874A1
GENERAL INFORMATION:
APPLICANT: Barney, S.
APPLICANT: Guthrie, K.
APPLICANT: Merutka, G.
APPLICANT: Anwer, M.
APPLICANT: Lambert, D.
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
FILE REFERENCE: 7872-100
CURRENT APPLICATION NUMBER: US/10/351,641
CURRENT FILING DATE: 2003-01-24
PRIOR APPLICATION NUMBER: 09/350,641
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/315,304
PRIOR FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/082,279
PRIOR FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 1757
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 553
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-10-351-641-553

Query Match 100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 LNNWFNI 7
Db 23 LNNWFNI 29

RESULT 9
US-10-351-641-554
Sequence 554, Application US/10351641
Publication No. US20030186874A1
GENERAL INFORMATION:
APPLICANT: Barney, S.
APPLICANT: Guthrie, K.
APPLICANT: Merutka, G.
APPLICANT: Anwer, M.
APPLICANT: Lambert, D.
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
FILE REFERENCE: 7872-100
CURRENT APPLICATION NUMBER: US/10/351,641
CURRENT FILING DATE: 2003-01-24
PRIOR APPLICATION NUMBER: 09/350,641
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/315,304
PRIOR FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/082,279
PRIOR FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 1757
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 554
LENGTH: 36
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-10-351-641-554

Query Match 100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 LNNWNI 7
 Db 24 LNNWNI 30

RESULT 10
 US-10-351-641-555
 ; Sequence 555, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.
 ; APPLICANT: Anwer, M.
 ; APPLICANT: Lambert, D.
 ; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
 ; FILE REFERENCE: 7872-100
 ; CURRENT APPLICATION NUMBER: US/10/351,641
 ; CURRENT FILING DATE: 2003-01-24
 ; PRIOR APPLICATION NUMBER: 09/350,641
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/315,304
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/082,279
 ; PRIOR FILING DATE: 1998-05-20
 ; NUMBER OF SEQ ID NOS: 1757
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 555
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Core polypeptide
 US-10-351-641-555

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWNI 7
 Db 25 LNNWNI 31

RESULT 11
 US-10-351-641-556
 ; Sequence 556, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.
 ; APPLICANT: Anwer, M.
 ; APPLICANT: Lambert, D.
 ; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
 ; FILE REFERENCE: 7872-100
 ; CURRENT APPLICATION NUMBER: US/10/351,641
 ; CURRENT FILING DATE: 2003-01-24
 ; PRIOR APPLICATION NUMBER: 09/350,641
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/315,304
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/082,279
 ; PRIOR FILING DATE: 1998-05-20
 ; NUMBER OF SEQ ID NOS: 1757
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 556
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence

; FEATURE:
 ; OTHER INFORMATION: Core polypeptide
 US-10-351-641-556

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWNI 7
 Db 26 LNNWNI 32

RESULT 12
 US-10-351-641-557
 ; Sequence 557, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.
 ; APPLICANT: Anwer, M.
 ; APPLICANT: Lambert, D.
 ; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
 ; FILE REFERENCE: 7872-100
 ; CURRENT APPLICATION NUMBER: US/10/351,641
 ; CURRENT FILING DATE: 2003-01-24
 ; PRIOR APPLICATION NUMBER: 09/350,641
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/315,304
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/082,279
 ; PRIOR FILING DATE: 1998-05-20
 ; NUMBER OF SEQ ID NOS: 1757
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 557
 ; LENGTH: 36
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Core polypeptide
 US-10-351-641-557

Query Match 100.0%; Score 48; DB 14; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWNI 7
 Db 27 LNNWNI 33

RESULT 13
 US-10-351-641-558
 ; Sequence 558, Application US/10351641
 ; Publication No. US20030186874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barney, S.
 ; APPLICANT: Guthrie, K.
 ; APPLICANT: Merutka, G.
 ; APPLICANT: Anwer, M.
 ; APPLICANT: Lambert, D.
 ; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
 ; FILE REFERENCE: 7872-100
 ; CURRENT APPLICATION NUMBER: US/10/351,641
 ; CURRENT FILING DATE: 2003-01-24
 ; PRIOR APPLICATION NUMBER: 09/350,641
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/315,304
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/082,279


```
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 558
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-558
```

```
Query Match          100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LWNWFNI 7
        |||||
Db      28 LWNWFNI 34
```

RESULT 14

```
US-10-351-641-559
; Sequence 559, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; PRIOR FILING DATE: 2003-01-24
; PRIOR FILING DATE: 1999-07-09
; PRIOR FILING DATE: 1999-07-09
; PRIOR FILING DATE: 1999-07-09
; PRIOR FILING DATE: 1999-05-20
; PRIOR FILING DATE: 1999-05-20
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 559
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-559
```

```
Query Match          100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LWNWFNI 7
        |||||
Db      29 LWNWFNI 35
```

RESULT 15

```
US-10-351-641-560
; Sequence 560, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
```

```
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 560
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-560
```

```
Query Match          100.0%; Score 48; DB 14; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LWNWFNI 7
        |||||
Db      30 LWNWFNI 36
```

```
Search completed: February 25, 2004, 16:04:26
Job time : 6.8481 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:39 ; Search time 1.41772 Seconds
(without alignments)
257.096 Million cell updates/sec

Title: US-09-877-606-7

Perfect score: 48

Sequence: 1 LWNWFNI 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 48 | 100.0 | 851 | 1 ENV_HV1B8 | P04582 human immun |
| 2 | 48 | 100.0 | 852 | 1 ENV_HV1B8 | P12488 human immun |
| 3 | 48 | 100.0 | 853 | 1 ENV_HV1B8 | P19551 human immun |
| 4 | 48 | 100.0 | 853 | 1 ENV_HV1B8 | P12487 human immun |
| 5 | 48 | 100.0 | 855 | 1 ENV_HV1B8 | P04580 human immun |
| 6 | 48 | 100.0 | 856 | 1 ENV_HV1B8 | P03375 human immun |
| 7 | 48 | 100.0 | 856 | 1 ENV_HV1B8 | P04578 human immun |
| 8 | 48 | 100.0 | 856 | 1 ENV_HV1B8 | P04624 human immun |
| 9 | 48 | 100.0 | 856 | 1 ENV_HV1B8 | P070626 human immun |
| 10 | 48 | 100.0 | 856 | 1 ENV_HV1B8 | P05878 human immun |
| 11 | 48 | 100.0 | 861 | 1 ENV_HV1B8 | P03377 human immun |
| 12 | 43 | 89.6 | 843 | 1 ENV_HV1B8 | P35961 human immun |
| 13 | 43 | 89.6 | 846 | 1 ENV_HV1B8 | P18799 human immun |
| 14 | 43 | 89.6 | 847 | 1 ENV_HV1B8 | P19550 human immun |
| 15 | 43 | 89.6 | 847 | 1 ENV_HV1B8 | P05880 human immun |
| 16 | 43 | 89.6 | 852 | 1 ENV_HV1B8 | P19549 human immun |
| 17 | 43 | 89.6 | 853 | 1 ENV_HV1B8 | P04581 human immun |
| 18 | 43 | 89.6 | 854 | 1 ENV_HV1B8 | P17281 chimpanzee |
| 19 | 43 | 89.6 | 855 | 1 ENV_HV1B8 | P03378 human immun |
| 20 | 43 | 89.6 | 856 | 1 ENV_HV1B8 | P05877 human immun |
| 21 | 43 | 89.6 | 856 | 1 ENV_HV1B8 | P18772 human immun |
| 22 | 43 | 89.6 | 856 | 1 ENV_HV1B8 | P05881 human immun |
| 23 | 43 | 89.6 | 859 | 1 ENV_HV1B8 | P04583 human immun |
| 24 | 43 | 89.6 | 861 | 1 ENV_HV1B8 | P31819 human immun |
| 25 | 43 | 89.6 | 863 | 1 ENV_HV1B8 | P05882 human immun |
| 26 | 43 | 89.6 | 865 | 1 ENV_HV1B8 | P04579 human immun |
| 27 | 42 | 87.5 | 848 | 1 ENV_HV1B8 | P20871 human immun |
| 28 | 42 | 87.5 | 856 | 1 ENV_HV1B8 | P03376 human immun |
| 29 | 42 | 87.5 | 867 | 1 ENV_HV1B8 | P12489 human immun |
| 30 | 39 | 81.2 | 193 | 1 SARI_CABEL | Q23445 caenorhabd |
| 31 | 39 | 81.2 | 241 | 1 CTR3_YEAST | Q06686 saccharomyc |
| 32 | 39 | 81.2 | 394 | 1 GATR_MOUSE | P23336 mus musculu |
| 33 | 39 | 81.2 | 574 | 1 CDAS_THEET | P29964 thermoanaer |

ALIGNMENTS

RESULT 1

ENV_HV1B8
ID ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=851111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L., Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; K02011; AAA44661.1; -
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HKG; 31-OCT-93.
DR PDB; 1Q03; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENVSBH8.
DR GlycoSuiteDB; P04582;
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.
FT SIGNAL 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 31 506 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 507 851 BY SIMILARITY.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 440 BY SIMILARITY.

34 39 81.2 1028 1 FDXG HAEIN P46448 haemophilus
35 38 79.2 562 1 SYR2 BACAA Q81r81 bacillus an
36 38 79.2 610 1 NUSM SQAAC Q92z44 squallus aca
37 38 79.2 855 1 ENV HV10Y P20888 human immun
38 37 77.1 625 1 TR11 MOUSE O35305 mus musculu
39 36 75.0 207 1 YE39 METJA Q58834 methanococc
40 36 75.0 245 1 TEV1 BPT4 P13299 bacterioph
41 36 75.0 366 1 YF59 STAAM Q99tt7 staphylococ
42 36 75.0 401 1 SYV HAEIN P43836 haemophilus
43 36 75.0 491 1 YIJ0 YEAST P40499 saccharomyc
44 36 75.0 785 1 ISP4 SCHPO P40900 schizosacch
45 35 72.9 173 1 CTR5 SCHPO Q9p7f9 schizosacch

```

FT DISULFID 385 413 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 100.0%; Score 48; DB 1; Length 851;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
Db 664 LWNWENI 670

RESULT 2
ENV_HV1BN STANDARD; PRT; 852 AA.
AC F12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (IBR isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P., Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency virus (HIV-1BR) from the brain of a patient with progressive dementia.";
RT Virology 168:79-89(1989).
CC -!- MISCELLANEOUS; THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

CC EMBL; M21098; AAA44221.1; -.
DR PIR; A31667; VCLJBR.
DR PDB; 1IM7; 23-OCT-02.
DR HIV; M21098; ENVSRVA.
DR InterPro; IPR00328; Env GP41.
DR InterPro; IPR00777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 507 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 508 852 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 155 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 236 330 BY SIMILARITY.
FT DISULFID 376 439 BY SIMILARITY.
FT DISULFID 383 412 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2B866345DEC915F CRC64;

Query Match 100.0%; Score 48; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
Db 665 LWNWENI 671

RESULT 3
ENV_HV1BN STANDARD; PRT; 853 AA.
AC P19551.
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.

```



```

FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;

Query Match 100.0%; Score 48; DB 1; Length 853;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
Db 666 LWNWENI 672

RESULT 5
ENV_HV126 STANDARD; PRT; 855 AA.
ID ENV_HV126 STANDARD; PRT; 855 AA.
AC P04560;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]

SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaram V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
Zaire: nucleotide sequence analysis identifies conserved and variable
domains in the envelope gene.";
RL Gene 52:71-82(1987).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03458; AAA45380.1; -.
DR PIR; D26192; VCLJ2R.
DR HIV; K03458; ENV526.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 511 855 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 207 BY SIMILARITY.
FT DISULFID 125 198 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 220 249 BY SIMILARITY.

```

```

FT DISULFID 230 241 BY SIMILARITY.
FT DISULFID 298 332 BY SIMILARITY.
FT DISULFID 378 444 BY SIMILARITY.
FT DISULFID 385 417 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 855 AA; 96971 MW; 3B4D3D6E239C3457 CRC64;

Query Match 100.0%; Score 48; DB 1; Length 855;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
Db 666 LWNWENI 674

RESULT 6
ENV_HV1B1 STANDARD; PRT; 856 AA.
ID ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doan E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Laubenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;

```



```
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AP85107FE0 CRC64;

Query Match 100.0%; Score 48; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 669 LMNWFNI 675
|||||

RESULT 8
ENV_HV1H3 STANDARD; PRT; 856 AA.
AC P04624;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein gp160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11707;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85228248; PubMed=2988795;
RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
RA Shaw G.M., Wong-Staal F., Reddy E.P.;
RT "HTLV-III env gene products synthesized in E. coli are recognized by
antibodies present in the sera of AIDS patients.";
RL Cell 41:979-986(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14100; AAA44679.1; -.
CC PDB; 1JAU; 17-OCT-01.
CC PDB; 1JAV; 17-OCT-01.
CC DR
CC DR
```

```
DR HIV; M14100; ENVSHXB3.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GPI20.
DR Pfam; PF00516; GPI20; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 31 511 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 512 905 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MW; 3373C68B8B4C1AFC CRC64;

Query Match 100.0%; Score 48; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 669 LMNWFNI 675
|||||

RESULT 9
ENV_HV1LW STANDARD; PRT; 856 AA.
ID ENV_HV1LW
AC Q70626;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein gp160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=82834;
OX
```



```

FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600EBA7A08 CRC64;

Query Match 100.0%; Score 48; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMNWFNI 7
Db 669 LMNWFNI 675
|||||

RESULT 11
ENV_HV1BR STANDARD; PRT; 861 AA.
AC F03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02013; AAB59751.1; -.
DR EMBL; A04321; CAA00352.1; -.
DR PIR; A03975; VCLJLV.
DR PDB; 1ERF; 20-FEB-02.
DR HIV; K02013; ENV$BRU.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure. 1 30
FT SIGNAL 1 30
FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 517 861 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 210 BY SIMILARITY.

```

```

FT DISULFID 126 201 BY SIMILARITY.
FT DISULFID 131 162 BY SIMILARITY.
FT DISULFID 223 252 BY SIMILARITY.
FT DISULFID 233 244 BY SIMILARITY.
FT DISULFID 301 336 BY SIMILARITY.
FT DISULFID 383 450 BY SIMILARITY.
FT DISULFID 390 423 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 642 642 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 861 AA; 97487 MW; 04DE2B4D4E4FD63A CRC64;

Query Match 100.0%; Score 48; DB 1; Length 861;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMNWFNI 7
Db 674 LMNWFNI 680
|||||

RESULT 12
ENV_HV1Y2 STANDARD; PRT; 843 AA.
ID ENV_HV1Y2
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -----

```



```

SQ SEQUENCE 846 AA; 96476 MW; 8A3B9DA527DE2B83 CRC64;
Query Match 89.6%; Score 43; DB 1; Length 846;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
|||||:|
Db 659 LWNWFSI 665

RESULT 14
ENV_HV1S1
ID ENV_HV1S1 STANDARD; PRT; 847 AA.
AC F19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M65024; AAA45072.1; -.
DR PDB; 1OBE; 15-MAY-97.
DR HIV; M38428; ENV$SF162.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 503 847 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 89.6%; Score 43; DB 1; Length 847;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
|||||:|
Db 660 LWNWFDI 666

RESULT 15
ENV_HV1W2
ID ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=862335450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
at risk for AIDS.";
RL Science 232:1548-1553(1986).
CC -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12507; AAB12990.1; -.
DR HIV; M12507; ENV$WMJ2.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW SIGNAL.
FT SIGNAL 1 29 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 30 501 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 502 847 BY SIMILARITY.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 202 BY SIMILARITY.
FT DISULFID 125 193 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 215 244 BY SIMILARITY.

```

| | | | | | |
|----|----------|---------|-----------|-------------------------|--------------|
| FT | DISULFID | 225 | 236 | BY SIMILARITY. | |
| FT | DISULFID | 293 | 326 | BY SIMILARITY. | |
| FT | DISULFID | 372 | 435 | BY SIMILARITY. | |
| FT | DISULFID | 379 | 408 | BY SIMILARITY. | |
| FT | CARBOHYD | 87 | 87 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 134 | 134 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 140 | 140 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 151 | 151 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 155 | 155 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 183 | 183 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 184 | 184 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 194 | 194 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 231 | 231 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 238 | 238 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 259 | 259 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 273 | 273 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 286 | 286 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 292 | 292 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 327 | 327 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 334 | 334 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 350 | 350 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 356 | 356 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 380 | 380 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 386 | 386 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 390 | 390 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 400 | 400 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 438 | 438 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 450 | 450 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 602 | 602 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 607 | 607 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 616 | 616 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| FT | CARBOHYD | 628 | 628 | N-LINKED (GLCNAC. . .) | (POTENTIAL). |
| SQ | SEQUENCE | 847 AA; | 96466 MW; | CD1E33D73AA5BCAE CRC64; | |

Query Match 89.6%; Score 43; DB 1; Length 847;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LWNWFI 7
| | | | |
Db 660 LWNWFI 666

Search completed: February 25, 2004, 15:41:48
Job time : 2.41772 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:35:54 ; Search time 6.86709 Seconds
(without alignments)
321.625 Million cell updates/sec

Title: US-09-877-606-7

Perfect score: 48

Sequence: 1 LNNWFNI 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_protein.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriophage.*
17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 48 | 100.0 | 117 | 15 | Q9YRT8 human immun |
| 2 | 48 | 100.0 | 117 | 15 | Q9YRT3 human immun |
| 3 | 48 | 100.0 | 117 | 15 | Q9YRT6 human immun |
| 4 | 48 | 100.0 | 117 | 15 | Q9YRS7 human immun |
| 5 | 48 | 100.0 | 117 | 15 | Q7SVI7 human immun |
| 6 | 48 | 100.0 | 117 | 15 | Q7SVH7 human immun |
| 7 | 48 | 100.0 | 117 | 15 | Q7SVF3 human immun |
| 8 | 48 | 100.0 | 118 | 15 | Q9E5R8 human immun |
| 9 | 48 | 100.0 | 118 | 15 | Q9E5S6 human immun |
| 10 | 48 | 100.0 | 118 | 15 | Q9E5P4 human immun |
| 11 | 48 | 100.0 | 118 | 15 | Q9E5P9 human immun |
| 12 | 48 | 100.0 | 118 | 15 | Q9E5P3 human immun |
| 13 | 48 | 100.0 | 118 | 15 | Q9E5Q8 human immun |
| 14 | 48 | 100.0 | 118 | 15 | Q9E5R9 human immun |
| 15 | 48 | 100.0 | 118 | 15 | Q9E5Q2 human immun |
| 16 | 48 | 100.0 | 121 | 15 | Q8UQX7 human immun |

| | | | | | | |
|----|----|-------|-----|----|--------|--------------------|
| 17 | 48 | 100.0 | 122 | 15 | Q9WR03 | Q9wr03 human immun |
| 18 | 48 | 100.0 | 122 | 15 | Q9OPY4 | Q9opy4 human immun |
| 19 | 48 | 100.0 | 122 | 15 | Q9QIW0 | Q9qiwo human immun |
| 20 | 48 | 100.0 | 122 | 15 | Q9EA96 | Q9ea96 human immun |
| 21 | 48 | 100.0 | 122 | 15 | Q9EAA0 | Q9eaa0 human immun |
| 22 | 48 | 100.0 | 122 | 15 | Q9ODK1 | Q9odk1 human immun |
| 23 | 48 | 100.0 | 122 | 15 | Q9QIW6 | Q9qiwo human immun |
| 24 | 48 | 100.0 | 122 | 15 | Q9YXR6 | Q9yxr6 human immun |
| 25 | 48 | 100.0 | 122 | 15 | Q9EAA5 | Q9eaa5 human immun |
| 26 | 48 | 100.0 | 122 | 15 | Q9QIU9 | Q9qiu9 human immun |
| 27 | 48 | 100.0 | 122 | 15 | Q9QDT0 | Q9qdt0 human immun |
| 28 | 48 | 100.0 | 122 | 15 | Q9YXN7 | Q9yxn7 human immun |
| 29 | 48 | 100.0 | 122 | 15 | Q9ILK1 | Q9ilk1 human immun |
| 30 | 48 | 100.0 | 122 | 15 | Q9YXN4 | Q9yxn4 human immun |
| 31 | 48 | 100.0 | 122 | 15 | Q9QIW9 | Q9qiwo human immun |
| 32 | 48 | 100.0 | 122 | 15 | Q9YXP1 | Q9yxp1 human immun |
| 33 | 48 | 100.0 | 122 | 15 | Q9ILQ7 | Q9ilq7 human immun |
| 34 | 48 | 100.0 | 122 | 15 | Q9YXN8 | Q9yxn8 human immun |
| 35 | 48 | 100.0 | 122 | 15 | Q9YXN8 | Q9yxn8 human immun |
| 36 | 48 | 100.0 | 122 | 15 | Q9QDS9 | Q9qds9 human immun |
| 37 | 48 | 100.0 | 122 | 15 | Q9ILK0 | Q9ilk0 human immun |
| 38 | 48 | 100.0 | 122 | 15 | Q9IJP2 | Q9ijp2 human immun |
| 39 | 48 | 100.0 | 122 | 15 | Q9ILK7 | Q9ilk7 human immun |
| 40 | 48 | 100.0 | 122 | 15 | Q9YXN1 | Q9yxn1 human immun |
| 41 | 48 | 100.0 | 122 | 15 | Q9IJQ1 | Q9ijq1 human immun |
| 42 | 48 | 100.0 | 122 | 15 | Q9IJL1 | Q9ijl1 human immun |
| 43 | 48 | 100.0 | 122 | 15 | Q9IJP9 | Q9ijp9 human immun |
| 44 | 48 | 100.0 | 122 | 15 | Q9QDP1 | Q9qdp1 human immun |
| 45 | 48 | 100.0 | 122 | 15 | Q9YXN5 | Q9yxn5 human immun |

ALIGNMENTS

RESULT 1

Q9YRT8 PRELIMINARY; PRT; 117 AA.
AC Q9YRT8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96USNG31;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.,
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096327; AAD04402.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0003198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
FT Transmembrane.
FT NON_TER 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 14156 MW; F3C6C7324386A240 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;

Best Local Similarity 100.0%; Pred. No. 2.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFNI 7

DB 109 LNNWFNI 115

```

RESULT 2
Q9YRT3 ID Q9YRT3 PRELIMINARY; PRT; 117 AA.
AC Q9YRT3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96USCm1;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.;
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096332; AAD04407.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14088 MW; 32ADA489B31BAFF4 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
DB 109 LWNWENI 115

RESULT 3
Q9YRK6 ID Q9YRK6 PRELIMINARY; PRT; 117 AA.
AC Q9YRK6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96USCm48;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.;
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096350; AAD04425.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14088 MW; 32ADA489B31BAFF4 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
DB 109 LWNWENI 115

RESULT 4
Q9YRS7 ID Q9YRS7 PRELIMINARY; PRT; 117 AA.
AC Q9YRS7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96USSN07;
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,
RA Jones J., Ward J.;
RT "Surveillance of Central African Nationals living in the United States
RT reveals multiple subtypes of HIV-1 Group M and Group O.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF096339; AAD04414.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14133 MW; ECBAF666A1CDA17D CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWENI 7
DB 109 LWNWENI 115

RESULT 5
Q7SVI7 ID Q7SVI7 PRELIMINARY; PRT; 117 AA.
AC Q7SVI7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein gp41 domain (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99NG1056;
RA Zeh C., Agwale S.M., Robbins K.E., Odama L., Sani-Gwarco N.,
RA Gboun M.S., Pieniazek D., Wambele C., Folks T.M., Kalish M.L.;
RT "Presence of HIV-2 subtype B in single and HIV-2 subtype A in dual
RT HIV-1/HIV-2 infections in Nigeria.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY102797; AAMS1892.1; -.

```



```

KW Envelope protein.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14054 MW; 77789BE3BA77775B CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
Db 109 LWNWFI 115

RESULT 6
Q7SVH7 PRELIMINARY; PRT; 117 AA.
AC Q7SVH7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein gp41 domain (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99NG7237;
RA Zhen C., Agwale S.M., Robbins K.E., Odama L., Sani-Gwarco N.,
RA Ghoun M.S., Pieniazek D., Wambebe C., Folks T.M., Kalish M.L.;
RT "Presence of HIV-2 subtype B in single and HIV-2 subtype A in dual
RT HIV-1/HIV-2 infections in Nigeria.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY102807; AA51902.1; -
KW Envelope protein.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 14064 MW; EC676FCA1B62CFA8 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
Db 109 LWNWFI 115

RESULT 7
Q7SVF3 PRELIMINARY; PRT; 117 AA.
AC Q7SVF3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein gp41 domain (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99NG77;
RA Zhen C., Agwale S.M., Robbins K.E., Odama L., Sani-Gwarco N.,
RA Ghoun M.S., Pieniazek D., Wambebe C., Folks T.M., Kalish M.L.;
RT "Presence of HIV-2 subtype B in single and HIV-2 subtype A in dual
RT HIV-1/HIV-2 infections in Nigeria.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY102831; AA51926.1; -
KW Envelope protein.
FT NON_TER 1
FT NON_TER 117

```

```

SQ SEQUENCE 117 AA; 14029 MW; 92B245BDB91467C6 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
Db 109 LWNWFI 115

RESULT 8
Q9ESR8 PRELIMINARY; PRT; 118 AA.
AC Q9ESR8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM161;
RX MEDLINE=20414627; PubMed=10957729;
RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,
RA John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral
RT Variants in Cameroon.";
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM161;
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemnji G.A., Eno L.T.,
RA Nguingsong J.N., Gao F., Rayfield M., Folks T.M., Pieniazek D.,
RA Lal R.B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP252103; AAG14312.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 14244 MW; CBE698BE4A3748C3 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
Db 105 LWNWFI 111

RESULT 9
Q9ESS6 PRELIMINARY; PRT; 118 AA.
AC Q9ESS6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=CAM103;
RX MEDLINE=20414627; PubMed=10957729;
RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,
RA John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral
RT Variants in Cameroon.";
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM103;
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemnji G.A., Eno L.T.,
RA Ngengasong J.N., Gao F., Rayfield M., Folks T.M., Pieniazek D.,
RA Lal R.B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252095; AAG14304.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 14257 MW; C11CF6BDEFBCB50 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
DB 105 LWNWFI 111

RESULT 10
Q9ESP4 PRELIMINARY; PRT; 118 AA.
AC Q9ESP4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM63;
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemnji G.A., Eno L.T.,
RA Ngengasong J.N., Gao F., Rayfield M., Folks T.M., Pieniazek D.,
RA Lal R.B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252127; AAG14336.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 14274 MW; 8B0ABECF5124B584 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
DB 105 LWNWFI 111

RESULT 11
Q9ESP9 PRELIMINARY; PRT; 118 AA.
AC Q9ESP9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM256;
RX MEDLINE=20414627; PubMed=10957729;
RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,
RA John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral
RT Variants in Cameroon.";
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM256;
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemnji G.A., Eno L.T.,
RA Ngengasong J.N., Gao F., Rayfield M., Folks T.M., Pieniazek D.,
RA Lal R.B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252122; AAG14331.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 14272 MW; 2C1E108925F4283C CRC64;

Query Match 100.0%; Score 48; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWNWFI 7
DB 105 LWNWFI 111

RESULT 12
Q9ESP3 PRELIMINARY; PRT; 118 AA.
AC Q9ESP3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM70;
RX MEDLINE=20414627; PubMed=10957729;

```

| | | | | | | | | | |
|--|--|---|----|------------|----|--------|-----|------|----|
| Matches | 7; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 1 | LWNWFNI 7 | | | | | | | |
| | | | | | | | | | |
| Db | 105 | LWNWFNI 111 | | | | | | | |
| RESULT 14 | | | | | | | | | |
| Q9E5R9 | | PRELIMINARY; | | PRT; | | 118 | AA. | | |
| ID | Q9E5R9 | | | | | | | | |
| AC | Q9E5R9; | | | | | | | | |
| DT | 01-MAR-2001 | (TrEMBLrel. 16, Created) | | | | | | | |
| DT | 01-MAR-2001 | (TrEMBLrel. 16, Last sequence update) | | | | | | | |
| DT | 01-JUN-2003 | (TrEMBLrel. 24, Last annotation update) | | | | | | | |
| DE | Envelope glycoprotein (Fragment). | | | | | | | | |
| DE | ENV. | | | | | | | | |
| OS | Human immunodeficiency virus 1. | | | | | | | | |
| OC | Viruses; Retroid viruses; Retroviridae; Lentivirus. | | | | | | | | |
| OX | NCBI_TaxID=11676; | | | | | | | | |
| RN | [1] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | STRAIN=CAM16; | | | | | | | | |
| RX | MEDLINE=20414627; PubMed=10957729; | | | | | | | | |
| RA | Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T., | | | | | | | | |
| RA | John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.; | | | | | | | | |
| RT | "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral | | | | | | | | |
| RT | Variants in Cameroon."; | | | | | | | | |
| RL | AIDS Res. Hum. Retroviruses 16:1319-1324(2000). | | | | | | | | |
| RN | [2] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | STRAIN=CAM16; | | | | | | | | |
| RA | Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemnji G.A., Eno L.T., | | | | | | | | |
| RA | Ngengasong J.N., Gao P., Rayfield M., Folke T.M., Pieniazek D., | | | | | | | | |
| RA | Lal R.B.; | | | | | | | | |
| RL | Submitted (APR-2000) to the EMBL/GenBank/DBJ databases. | | | | | | | | |
| DR | EMBL; AF252102; AAG14311.1; -. | | | | | | | | |
| DR | GO; GO:0016021; C:integral to membrane; IEA. | | | | | | | | |
| DR | GO; GO:0019031; C:viral envelope; IEA. | | | | | | | | |
| DR | GO; GO:0005198; F:structural molecule activity; IEA. | | | | | | | | |
| DR | InterPro: IPR000328; Env_GP41. | | | | | | | | |
| DR | Pfam; PF00517; GP41; 1. | | | | | | | | |
| KW | Transmembrane. | | | | | | | | |
| FT | NON_TER 1 | | | | | | | | |
| FT | NON_TER 118 118 | | | | | | | | |
| SQ | SEQUENCE 118 AA; 14159 MW; 54739ADDD14A51F7 CRC64; | | | | | | | | |
| Query Match 100.0%; Score 48; DB 15; Length 118; | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 2.2; | | | | | | | | | |
| Matches | 7; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 1 | LWNWFNI 7 | | | | | | | |
| | | | | | | | | | |
| Db | 105 | LWNWFNI 111 | | | | | | | |
| RESULT 15 | | | | | | | | | |
| Q9E5Q2 | | PRELIMINARY; | | PRT; | | 118 | AA. | | |
| ID | Q9E5Q2 | | | | | | | | |
| AC | Q9E5Q2; | | | | | | | | |
| DT | 01-MAR-2001 | (TrEMBLrel. 16, Created) | | | | | | | |
| DT | 01-MAR-2001 | (TrEMBLrel. 16, Last sequence update) | | | | | | | |
| DT | 01-JUN-2003 | (TrEMBLrel. 24, Last annotation update) | | | | | | | |
| DE | Envelope glycoprotein (Fragment). | | | | | | | | |
| DE | ENV. | | | | | | | | |
| OS | Human immunodeficiency virus 1. | | | | | | | | |
| OC | Viruses; Retroid viruses; Retroviridae; Lentivirus. | | | | | | | | |
| OX | NCBI_TaxID=11676; | | | | | | | | |
| RN | [1] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | STRAIN=CAM232; | | | | | | | | |
| RX | MEDLINE=20414627; PubMed=10957729; | | | | | | | | |
| RA | Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T., | | | | | | | | |
| RA | John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.; | | | | | | | | |

```

RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral
RL Variants in Cameroon.";
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CAM232;
RA Fonjungo P.N., Mpoudi E.N., Torimiro J.N., Alemnji G.A., Eno L.T.,
RA Ngengeasong J.N., Gao F., Rayfield M., Folks T.M., Pieniazek D.,
RA Lal R.B.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252119; AAG14328.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 14295 MW; 807EDE44F29796AC CRC64;

Query Match 100.0%; Score 48; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. NO. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNNWFNI 7
Db 105 LNNWFNI 111

```

Search completed: February 25, 2004, 15:44:41
Job time : 8.86709 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:09 ; Search time 10.3228 Seconds
(without alignments)
191.599 Million cell updates/sec

Title: US-09-877-606-8

Perfect score: 7

Sequence: 1 XXXXXX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 0 | 0.0 | 1 | 2 | ADD95004 Platelet |
| 2 | 0 | 0.0 | 1 | 2 | ADD94992 Platelet |
| 3 | 0 | 0.0 | 1 | 2 | ADD94993 Platelet |
| 4 | 0 | 0.0 | 1 | 2 | ADD95002 Platelet |
| 5 | 0 | 0.0 | 1 | 2 | ADD94997 Platelet |
| 6 | 0 | 0.0 | 1 | 2 | ADD94999 Platelet |
| 7 | 0 | 0.0 | 1 | 2 | ADD95003 Platelet |
| 8 | 0 | 0.0 | 1 | 2 | ADD94995 Platelet |
| 9 | 0 | 0.0 | 1 | 2 | ADD94998 Platelet |
| 10 | 0 | 0.0 | 1 | 2 | ADD95001 Platelet |
| 11 | 0 | 0.0 | 1 | 2 | ADD94990 Platelet |
| 12 | 0 | 0.0 | 1 | 2 | ADD94991 Platelet |
| 13 | 0 | 0.0 | 1 | 2 | ADD94994 Platelet |
| 14 | 0 | 0.0 | 1 | 2 | ADD94996 Platelet |
| 15 | 0 | 0.0 | 1 | 2 | ADD95000 Platelet |
| 16 | 0 | 0.0 | 1 | 2 | RAY46652 Immunogen |
| 17 | 0 | 0.0 | 1 | 4 | AAM97834 Human pep |
| 18 | 0 | 0.0 | 1 | 4 | AAM97974 Human pep |
| 19 | 0 | 0.0 | 1 | 4 | AAM97643 Human pep |
| 20 | 0 | 0.0 | 1 | 4 | AAM98447 Human pep |
| 21 | 0 | 0.0 | 1 | 4 | AAM98354 Human pep |
| 22 | 0 | 0.0 | 1 | 4 | AAM93290 Human non |
| 23 | 0 | 0.0 | 1 | 4 | AAM53329 Human non |
| 24 | 0 | 0.0 | 1 | 4 | AAM53291 Human non |
| 25 | 0 | 0.0 | 1 | 4 | AAM53328 Human non |

| | | | | | |
|----|---|-----|---|---|--------------------|
| 26 | 0 | 0.0 | 1 | 4 | AAM53219 Human non |
| 27 | 0 | 0.0 | 1 | 4 | AAM53218 Human non |
| 28 | 0 | 0.0 | 1 | 4 | AAB91029 Thyrotrop |
| 29 | 0 | 0.0 | 1 | 4 | AAB91739 Opioid pe |
| 30 | 0 | 0.0 | 1 | 4 | AAB92150 Polypepti |
| 31 | 0 | 0.0 | 1 | 4 | AAB91892 Apoptosis |
| 32 | 0 | 0.0 | 1 | 4 | AAB91546 Endotheli |
| 33 | 0 | 0.0 | 1 | 4 | AAB92392 Miscellan |
| 34 | 0 | 0.0 | 1 | 4 | AAB91665 Opioid pe |
| 35 | 0 | 0.0 | 1 | 4 | AAG99966 ERA bindi |
| 36 | 0 | 0.0 | 1 | 4 | AAM00011 ERA bindi |
| 37 | 0 | 0.0 | 1 | 4 | AAG99987 ERA bindi |
| 38 | 0 | 0.0 | 1 | 4 | AAG99983 ERA bindi |
| 39 | 0 | 0.0 | 1 | 4 | AAM00013 ERA bindi |
| 40 | 0 | 0.0 | 1 | 4 | AAM00016 ERA bindi |
| 41 | 0 | 0.0 | 1 | 4 | AAG99988 ERA bindi |
| 42 | 0 | 0.0 | 1 | 4 | AAM00010 ERA bindi |
| 43 | 0 | 0.0 | 1 | 4 | ABB66809 Drosophil |
| 44 | 0 | 0.0 | 1 | 4 | ABB66810 Drosophil |
| 45 | 0 | 0.0 | 1 | 4 | AAG98134 Human SNP |

ALIGNMENTS

RESULT 1
ADD95004
ID ADD95004 standard; peptide; 1 AA.
XX AC ADD95004;
XX 29-JAN-2004 (first entry)
XX Platelet aggregation inhibitor peptide #146.
XX platelet aggregation inhibitor; guanidino group; amidino group.
XX Unidentified.

Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"

WO9501371-Al.

12-JAN-1995.

22-JUN-1994; 94WO-JP000999.

30-JUN-1993; 93JP-00186755.

(YAWA) NIPPON STEEL CORP.

Sato Y, Hayashi Y, Katada J, Takiguchi Y;

WPI; 1995-060950/08.

New RGD peptide(s) useful as anti:platelet aggregation agents - contain guanidino or amidino gp. at N-terminal to increase stability.

Disclosure; Page 11; 34pp; Japanese.

The invention describes peptides of amino acid sequence (I) and their salts. (I) are useful as platelet aggregation inhibitors and are easily absorbed by the body. Due to the presence of the N-terminal guanidino or amidino group, (I) are provided with excellent stability so that their activity can be exhibited for an effective time after administration. Thereafter they are readily metabolised and expelled. This is the amino acid sequence of a platelet aggregation inhibitor peptide.

Sequence 1 AA;

```

Query Match      0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 2
ADD94992
ID ADD94992 standard; peptide; 1 AA.
XX
AC ADD94992;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #134.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
FT defined)"
XX
PN WO9501371-Al.
XX
PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAMA ) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
WPI; 1995-060950/08.
XX
New RGD peptide(s) useful as anti:platelet aggregation agents - contain
guanidino or amidino gp. at N-terminal to increase stability.
XX
PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAMA ) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
WPI; 1995-060950/08.
XX
New RGD peptide(s) useful as anti:platelet aggregation agents - contain
guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
salts. (I) are useful as platelet aggregation inhibitors and are easily
absorbed by the body. Due to the presence of the N-terminal guanidino or
amidino group, (I) are provided with excellent stability so that their
activity can be exhibited for an effective time after administration.
Thereafter they are readily metabolised and expelled. This is the amino
acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;

Query Match      0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 3
ADD94993
ID ADD94993 standard; peptide; 1 AA.
XX
AC ADD94993;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #144.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
FT defined)"
XX
PN WO9501371-Al.
XX
PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAMA ) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
WPI; 1995-060950/08.
XX
New RGD peptide(s) useful as anti:platelet aggregation agents - contain
guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
salts. (I) are useful as platelet aggregation inhibitors and are easily
absorbed by the body. Due to the presence of the N-terminal guanidino or
amidino group, (I) are provided with excellent stability so that their
activity can be exhibited for an effective time after administration.
Thereafter they are readily metabolised and expelled. This is the amino
acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;

```

```

XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #135.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
FT defined)"
XX
PN WO9501371-Al.
XX
PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAMA ) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
WPI; 1995-060950/08.
XX
New RGD peptide(s) useful as anti:platelet aggregation agents - contain
guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
salts. (I) are useful as platelet aggregation inhibitors and are easily
absorbed by the body. Due to the presence of the N-terminal guanidino or
amidino group, (I) are provided with excellent stability so that their
activity can be exhibited for an effective time after administration.
Thereafter they are readily metabolised and expelled. This is the amino
acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;

Query Match      0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 4
ADD95002
ID ADD95002 standard; peptide; 1 AA.
XX
AC ADD95002;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #144.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
FT defined)"
XX

```

PN WO9501371-A1.
 XX 12-JAN-1995.
 XX
 XX 22-JUN-1994; 94WO-JP000999.
 XX 30-JUN-1993; 93JP-00186755.
 XX (YAMA) NIPPON STEEL CORP.
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;
 XX WPI; 1995-060950/08.
 XX
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain
 XX guanidino or amidino gp. at N-terminal to increase stability.
 XX Disclosure; Page 10-11; 34pp; Japanese.
 XX
 XX The invention describes peptides of amino acid sequence (I) and their
 XX salts. (I) are useful as platelet aggregation inhibitors and are easily
 XX absorbed by the body. Due to the presence of the N-terminal guanidino or
 XX amidino group, (I) are provided with excellent stability so that their
 XX activity can be exhibited for an effective time after administration.
 XX Thereafter they are readily metabolised and expelled. This is the amino
 XX acid sequence of a platelet aggregation inhibitor peptide.
 XX Sequence 1 AA;
 XX
 XX Query Match 0.0%; Score 0; DB 2; Length 1;
 XX Best Local Similarity 0.0%; Pred. No. 0;
 XX Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 1 R 1
 RESULT 5
 ADD94997
 ID ADD94997 standard; peptide; 1 AA.
 XX
 XX AC ADD94997;
 XX
 XX 29-JAN-2004 (first entry)
 XX Platelet aggregation inhibitor peptide #139.
 XX platelet aggregation inhibitor; guanidino group; amidino group.
 XX Unidentified.
 XX
 XX Key Location/Qualifiers
 XX Modified-site 1
 XX /label= OTHER
 XX /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
 XX defined)"
 XX
 XX WO9501371-A1.
 XX 12-JAN-1995.
 XX
 XX 22-JUN-1994; 94WO-JP000999.
 XX 30-JUN-1993; 93JP-00186755.
 XX (YAMA) NIPPON STEEL CORP.
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;
 XX WPI; 1995-060950/08.
 XX
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain

PT guanidino or amidino gp. at N-terminal to increase stability.
 XX Disclosure; Page 10; 34pp; Japanese.
 XX
 XX The invention describes peptides of amino acid sequence (I) and their
 XX salts. (I) are useful as platelet aggregation inhibitors and are easily
 XX absorbed by the body. Due to the presence of the N-terminal guanidino or
 XX amidino group, (I) are provided with excellent stability so that their
 XX activity can be exhibited for an effective time after administration.
 XX Thereafter they are readily metabolised and expelled. This is the amino
 XX acid sequence of a platelet aggregation inhibitor peptide.
 XX Sequence 1 AA;
 XX
 XX Query Match 0.0%; Score 0; DB 2; Length 1;
 XX Best Local Similarity 0.0%; Pred. No. 0;
 XX Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 1 R 1
 RESULT 6
 ADD94999
 ID ADD94999 standard; peptide; 1 AA.
 XX
 XX AC ADD94999;
 XX
 XX 29-JAN-2004 (first entry)
 XX Platelet aggregation inhibitor peptide #141.
 XX platelet aggregation inhibitor; guanidino group; amidino group.
 XX Unidentified.
 XX
 XX Key Location/Qualifiers
 XX Modified-site 1
 XX /label= OTHER
 XX /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
 XX defined)"
 XX
 XX WO9501371-A1.
 XX 12-JAN-1995.
 XX
 XX 22-JUN-1994; 94WO-JP000999.
 XX 30-JUN-1993; 93JP-00186755.
 XX (YAMA) NIPPON STEEL CORP.
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;
 XX WPI; 1995-060950/08.
 XX
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain
 XX guanidino or amidino gp. at N-terminal to increase stability.
 XX Disclosure; Page 10; 34pp; Japanese.
 XX
 XX The invention describes peptides of amino acid sequence (I) and their
 XX salts. (I) are useful as platelet aggregation inhibitors and are easily
 XX absorbed by the body. Due to the presence of the N-terminal guanidino or
 XX amidino group, (I) are provided with excellent stability so that their
 XX activity can be exhibited for an effective time after administration.
 XX Thereafter they are readily metabolised and expelled. This is the amino
 XX acid sequence of a platelet aggregation inhibitor peptide.
 XX Sequence 1 AA;
 XX
 XX Query Match 0.0%; Score 0; DB 2; Length 1;

PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAWA) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
DR WPI; 1995-060950/08.
XX
PT New RGD peptide(s) useful as anti:platelet aggregation agents - contain
PT guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
CC salts. (I) are useful as platelet aggregation inhibitors and are easily
CC absorbed by the body. Due to the presence of the N-terminal guanidino or
CC amidino group, (I) are provided with excellent stability so that their
CC activity can be exhibited for an effective time after administration.
CC Thereafter they are readily metabolised and expelled. This is the amino
CC acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 R 1
RESULT 10
ADD95001
ID ADD95001 standard; peptide; 1 AA.
XX
AC ADD95001;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #143.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
FT defined)"
XX
PN WO9501371-A1.
XX
PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAWA) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
DR WPI; 1995-060950/08.
XX
PT New RGD peptide(s) useful as anti:platelet aggregation agents - contain
PT guanidino or amidino gp. at N-terminal to increase stability.
XX

PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
CC salts. (I) are useful as platelet aggregation inhibitors and are easily
CC absorbed by the body. Due to the presence of the N-terminal guanidino or
CC amidino group, (I) are provided with excellent stability so that their
CC activity can be exhibited for an effective time after administration.
CC Thereafter they are readily metabolised and expelled. This is the amino
CC acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 R 1
RESULT 11
ADD94990
ID ADD94990 standard; peptide; 1 AA.
XX
AC ADD94990;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #132.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
FT defined)"
XX
PN WO9501371-A1.
XX
PD 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAWA) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX
DR WPI; 1995-060950/08.
XX
PT New RGD peptide(s) useful as anti:platelet aggregation agents - contain
PT guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
CC salts. (I) are useful as platelet aggregation inhibitors and are easily
CC absorbed by the body. Due to the presence of the N-terminal guanidino or
CC amidino group, (I) are provided with excellent stability so that their
CC activity can be exhibited for an effective time after administration.
CC Thereafter they are readily metabolised and expelled. This is the amino
CC acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 X 1
DB      1 R 1

RESULT 12
ADD94991
ID ADD94991 standard; peptide; 1 AA.
XX
AC ADD94991;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #133.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
XX defined)"
XX
PN WO9501371-Al.
XX
DT 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
XX defined)"
XX
PN WO9501371-Al.
XX
DT 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAWA ) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX WPI; 1995-060950/08.
XX
DR New RGD peptide(s) useful as anti:platelet aggregation agents - contain
PT guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
CC salts. (I) are useful as platelet aggregation inhibitors and are easily
CC absorbed by the body. Due to the presence of the N-terminal guanidino or
CC amidino group, (I) are provided with excellent stability so that their
CC activity can be exhibited for an effective time after administration.
CC Thereafter they are readily metabolised and expelled. This is the amino
CC acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;
XX
Query Match 0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 X 1
DB      1 R 1

RESULT 13
ADD94994
ID ADD94994 standard; peptide; 1 AA.
XX
AC ADD94994;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #136.
XX

```

```

XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
XX defined)"
XX
PN WO9501371-Al.
XX
DT 12-JAN-1995.
XX
PF 22-JUN-1994; 94WO-JP000999.
XX
PR 30-JUN-1993; 93JP-00186755.
XX
PA (YAWA ) NIPPON STEEL CORP.
XX
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX WPI; 1995-060950/08.
XX
DR New RGD peptide(s) useful as anti:platelet aggregation agents - contain
PT guanidino or amidino gp. at N-terminal to increase stability.
XX
PS Disclosure; Page 10; 34pp; Japanese.
XX
CC The invention describes peptides of amino acid sequence (I) and their
CC salts. (I) are useful as platelet aggregation inhibitors and are easily
CC absorbed by the body. Due to the presence of the N-terminal guanidino or
CC amidino group, (I) are provided with excellent stability so that their
CC activity can be exhibited for an effective time after administration.
CC Thereafter they are readily metabolised and expelled. This is the amino
CC acid sequence of a platelet aggregation inhibitor peptide.
XX
SQ Sequence 1 AA;
XX
Query Match 0.0%; Score 0; DB 2; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 X 1
DB      1 R 1

RESULT 14
ADD94996
ID ADD94996 standard; peptide; 1 AA.
XX
AC ADD94996;
XX
DT 29-JAN-2004 (first entry)
XX
DE Platelet aggregation inhibitor peptide #138.
XX
KW platelet aggregation inhibitor; guanidino group; amidino group.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
XX defined)"
XX
PN WO9501371-Al.
XX
DT 12-JAN-1995.
XX

```

PF 22-JUN-1994; 94WO-JP000999.
 XX 30-JUN-1993; 93JP-00186755.
 XX (YAWA) NIPPON STEEL CORP.
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;
 XX WPI; 1995-060950/08.
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain
 PT guanidino or amidino gp. at N-terminal to increase stability.
 XX Disclosure; Page 10; 34pp; Japanese.
 XX The invention describes peptides of amino acid sequence (I) and their
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or
 CC amidino group, (I) are provided with excellent stability so that their
 CC activity can be exhibited for an effective time after administration.
 CC Thereafter they are readily metabolised and expelled. This is the amino
 CC acid sequence of a platelet aggregation inhibitor peptide.
 XX Sequence 1 AA;
 SQ

Query Match 0.0%; Score 0; DB 2; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 R 1

RESULT 15
 ADD95000
 ID ADD95000 standard; peptide; 1 AA.
 XX AC ADD95000;
 XX DT 29-JAN-2004 (first entry)
 XX DE Platelet aggregation inhibitor peptide #142.
 XX KW platelet aggregation inhibitor; guanidino group; amidino group.
 XX OS Unidentified.
 XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /label= OTHER
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
 FT defined)"
 XX WO9501371-A1.
 XX 12-JAN-1995.
 XX 22-JUN-1994; 94WO-JP000999.
 XX 30-JUN-1993; 93JP-00186755.
 XX (YAWA) NIPPON STEEL CORP.
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;
 XX WPI; 1995-060950/08.
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain
 PT guanidino or amidino gp. at N-terminal to increase stability.
 XX Disclosure; Page 10; 34pp; Japanese.
 XX

CC The invention describes peptides of amino acid sequence (I) and their
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or
 CC amidino group, (I) are provided with excellent stability so that their
 CC activity can be exhibited for an effective time after administration.
 CC Thereafter they are readily metabolised and expelled. This is the amino
 CC acid sequence of a platelet aggregation inhibitor peptide.
 XX Sequence 1 AA;
 SQ

Query Match 0.0%; Score 0; DB 2; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 R 1

Search completed: February 25, 2004, 15:41:10
 Job time : 12.3228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:37:04 ; Search time 2.79114 Seconds
(without alignments)
129.475 Million cell updates/sec

Title: US-09-877-606-8

Perfect score: 7 XXXXXX 7

Sequence: 1 XXXXXX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*

2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*

3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*

4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*

5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*

6: /cgn2_6/ptodata/2/iaa/backfilei.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 0 | 0.0 | 1 | 1 | US-07-820-154A-12 |
| 2 | 0 | 0.0 | 1 | 1 | US-07-791-213D-24 |
| 3 | 0 | 0.0 | 1 | 1 | US-07-791-213D-40 |
| 4 | 0 | 0.0 | 1 | 1 | US-08-174-365A-57 |
| 5 | 0 | 0.0 | 1 | 1 | US-07-789-913-23 |
| 6 | 0 | 0.0 | 1 | 1 | US-07-789-913-25 |
| 7 | 0 | 0.0 | 1 | 1 | US-08-049-794-23 |
| 8 | 0 | 0.0 | 1 | 1 | US-08-049-794-25 |
| 9 | 0 | 0.0 | 1 | 1 | US-08-433-037-12 |
| 10 | 0 | 0.0 | 1 | 1 | US-08-448-606-4 |
| 11 | 0 | 0.0 | 1 | 1 | US-07-869-933-16 |
| 12 | 0 | 0.0 | 1 | 1 | US-08-293-150A-24 |
| 13 | 0 | 0.0 | 1 | 1 | US-08-293-150A-40 |
| 14 | 0 | 0.0 | 1 | 1 | US-08-496-847-23 |
| 15 | 0 | 0.0 | 1 | 1 | US-08-496-847-25 |
| 16 | 0 | 0.0 | 1 | 2 | US-08-742-774-23 |
| 17 | 0 | 0.0 | 1 | 2 | US-08-742-774-25 |
| 18 | 0 | 0.0 | 1 | 2 | US-08-675-354-23 |
| 19 | 0 | 0.0 | 1 | 2 | US-08-675-354-25 |
| 20 | 0 | 0.0 | 1 | 2 | US-08-097-554A-12 |
| 21 | 0 | 0.0 | 1 | 2 | US-08-965-918-23 |
| 22 | 0 | 0.0 | 1 | 2 | US-08-965-918-25 |
| 23 | 0 | 0.0 | 1 | 2 | US-09-138-439-23 |
| 24 | 0 | 0.0 | 1 | 2 | US-09-138-439-25 |
| 25 | 0 | 0.0 | 1 | 3 | US-08-480-640A-12 |
| 26 | 0 | 0.0 | 1 | 3 | US-08-613-400A-23 |
| 27 | 0 | 0.0 | 1 | 3 | US-08-613-400A-25 |

28 0 0.0 1 3 US-08-801-092-10 Sequence 10, Appl
29 0 0.0 1 3 US-08-801-092-17 Sequence 17, Appl
30 0 0.0 1 3 US-08-801-092-24 Sequence 24, Appl
31 0 0.0 1 3 US-08-801-092-31 Sequence 31, Appl
32 0 0.0 1 3 US-08-801-092-38 Sequence 38, Appl
33 0 0.0 1 3 US-08-801-092-45 Sequence 45, Appl
34 0 0.0 1 3 US-09-298-017-23 Sequence 23, Appl
35 0 0.0 1 3 US-09-298-017-25 Sequence 25, Appl
36 0 0.0 1 3 US-08-295-802-12 Sequence 12, Appl
37 0 0.0 1 3 US-09-392-979A-23 Sequence 23, Appl
38 0 0.0 1 3 US-09-392-979A-25 Sequence 25, Appl
39 0 0.0 1 3 US-09-103-663-16 Sequence 16, Appl
40 0 0.0 1 3 US-08-488-237A-12 Sequence 12, Appl
41 0 0.0 1 3 US-09-117-927-5 Sequence 5, Appl
42 0 0.0 1 4 US-08-375-992A-12 Sequence 12, Appl
43 0 0.0 1 4 US-09-315-113-10 Sequence 10, Appl
44 0 0.0 1 4 US-09-315-113-17 Sequence 17, Appl
45 0 0.0 1 4 US-09-315-113-24 Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-07-820-154A-12
; Sequence 12, Application US/07820154A
; Patent No. 5382425
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-820-154A-12

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 V 1

RESULT 2

US-07-791-213D-24

; Sequence 24, Application US/07791213D
; Patent No. 5409895
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,213D
; FILING DATE: 13-NOV-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 3
US-07-791-213D-40
; Sequence 40, Application US/07791213D
; Patent No. 5409895
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States

; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,213D
; FILING DATE: 13-NOV-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 4
US-08-174-365A-57
; Sequence 57, Application US/08174365A
; Patent No. 5478809
; GENERAL INFORMATION:
; APPLICANT: Seiichi TANIDA et al.
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,365A
; FILING DATE: December 28, 1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 57:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1 amino acid
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: modified site
; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: /note = "Xaa is modified amino acid as
; OTHER INFORMATION: described in specification"
; US-08-174-365A-57

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1

Db 1 X 1

RESULT 5

US-07-789-913-23
; Sequence 23, Application US/07789913

; Patent No. 5559095

; GENERAL INFORMATION:

; APPLICANT: Miljanich, George P.

; APPLICANT: Bowersox, Stephen S.

; APPLICANT: Fox, James A.

; APPLICANT: Valentino, Karen L.

; APPLICANT: Bitner, Robert S.

; APPLICANT: Yamashiro, Donald H.

; TITLE OF INVENTION: Delayed Treatment Method of Reducing

; TITLE OF INVENTION: Ischemia-Related Neuronal Damage

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Law Offices of Peter Dehlinger

; STREET: 350 Cambridge Avenue, Suite 300

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/789,913

; FILING DATE: 19911112

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/561,766

; FILING DATE: 02-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/440,094

; FILING DATE: 22-NOV-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Stratford, Carol A.

; REGISTRATION NUMBER: 34,444

; REFERENCE/DOCKET NUMBER: 5865-0005.30

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: both

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: peptide fragment used in the claims
; US-07-789-913-23

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 C 1

RESULT 6

US-07-789-913-25

; Sequence 25, Application US/07789913

; Patent No. 5559095

; GENERAL INFORMATION:

; APPLICANT: Miljanich, George P.

; APPLICANT: Bowersox, Stephen S.

; APPLICANT: Fox, James A.

; APPLICANT: Valentino, Karen L.

; APPLICANT: Bitner, Robert S.

; APPLICANT: Yamashiro, Donald H.

; TITLE OF INVENTION: Delayed Treatment Method of Reducing

; TITLE OF INVENTION: Ischemia-Related Neuronal Damage

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Law Offices of Peter Dehlinger

; STREET: 350 Cambridge Avenue, Suite 300

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/789,913

; FILING DATE: 19911112

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/561,766

; FILING DATE: 02-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/440,094

; FILING DATE: 22-NOV-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Stratford, Carol A.

; REGISTRATION NUMBER: 34,444

; REFERENCE/DOCKET NUMBER: 5865-0005.30

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: both

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: peptide fragment used in the claims

; US-07-789-913-25

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 R 1

RESULT 7

US-08-049-794-23
; Sequence 23, Application US/08049794
; Patent No. 5587454

GENERAL INFORMATION:

APPLICANT: JUSTICE, ALAN
APPLICANT: SINGH, TEJINDER
APPLICANT: GOHIL, KISHOR C
APPLICANT: VALENTINO, KAREN L
APPLICANT: MILJANICH, GEORGE P
TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND
TITLE OF INVENTION: ENHANCING OPIATE ANALGESIA
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Dehlinger
STREET: 350 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/049,794
FILING DATE: 19930415
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/814,759
FILING DATE: 30-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0009.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE
INDIVIDUAL ISOLATE: 32

US-08-049-794-23

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 C 1

RESULT 8

US-08-049-794-25
; Sequence 25, Application US/08049794
; Patent No. 5587454

GENERAL INFORMATION:

APPLICANT: JUSTICE, ALAN
APPLICANT: SINGH, TEJINDER
APPLICANT: GOHIL, KISHOR C
APPLICANT: VALENTINO, KAREN L

APPLICANT: MILJANICH, GEORGE P
TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND
TITLE OF INVENTION: ENHANCING OPIATE ANALGESIA
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Dehlinger
STREET: 350 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/049,794
FILING DATE: 19930415
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/814,759
FILING DATE: 30-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0009.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE
INDIVIDUAL ISOLATE: 32

US-08-049-794-25

Query Match

Best Local Similarity 0.0%; Score 0; DB 1; Length 1;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 R 1

RESULT 9

US-08-433-037-12
; Sequence 12, Application US/08433037
; Patent No. 5707828

GENERAL INFORMATION:

APPLICANT: Sreekrishna, Kotikanyadan
APPLICANT: Barr, Kathryn A.
APPLICANT: Briertley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Tschopp, Juerg F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PICHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/433,037
;; FILING DATE: 03-MAY-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Digiglio, Frank S.
;; REGISTRATION NUMBER: 31,346
;; REFERENCE/DOCKET NUMBER: 91082
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 742-4343
;; TELEFAX: (516) 742-4366
;; TELEX: 230 901 SANS UR
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-433-037-12

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 L 1

RESULT 10
US-08-448-606-4
; Sequence 4, Application US/08448606
; Patent No. 5721114
; GENERAL INFORMATION:
; APPLICANT: Abrahams n, Lars
; APPLICANT: Holmgren, Erik
; APPLICANT: Kalder n, Christina
; APPLICANT: Lake, Mats
; APPLICANT: Mikaelsson, sa
; APPLICANT: Sellitz, Torsten
; TITLE OF INVENTION: Expression System For Producing
; TITLE OF INVENTION: Apolipoprotein AI-M
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pollock, Vande Sande & Priddy
; STREET: 1990 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,606
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SE93/01061
; FILING DATE: 09-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9203753-0
; FILING DATE: 11-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Amernick, Burton A.
; REGISTRATION NUMBER: 24,852
; REFERENCE/DOCKET NUMBER: 0151/00121
; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (202)331-7111
;; TELEFAX: (202) 293-6229
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1 amino acid
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-448-606-4

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 11
US-07-869-933-16
; Sequence 16, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-869-933-16

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 M 1

RESULT 12
US-08-293-150A-24
; Sequence 24, Application US/08293150A

```

; Patent No. 5792629
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,150A
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 07/791,213
; FILING DATE: 13-NOV-1990
; APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-293-150A-24

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 13
US-08-293-150A-40
; Sequence 40, Application US/08293150A
; Patent No. 5792629
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria

```

```

; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,150A
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 07/791,213
; FILING DATE: 13-NOV-1990
; APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-293-150A-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 14
US-08-496-847-23
; Sequence 23, Application US/08496847
; Patent No. 5795864
; GENERAL INFORMATION:
; APPLICANT: Amstutz, Gary A.
; APPLICANT: Bowersox, Stephen S.
; APPLICANT: Gohil, Kishorchandra
; APPLICANT: Adriaenssens, Peter I.
; APPLICANT: Kristipati, Ramasharma
; TITLE OF INVENTION: METHODS AND
; FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306-1546
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,847
; FILING DATE: 27-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stratford, Carol A

```

REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0009.31
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT
US-08-496-847-23

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 15
US-08-496-847-25
Sequence 25, Application US/08496847
Patent No. 5795864
GENERAL INFORMATION:
APPLICANT: Amstutz, Gary A.
APPLICANT: Bowersox, Stephen S.
APPLICANT: Gohil, Kishorchandra
APPLICANT: Adriaenssens, Peter I.
APPLICANT: Kristipati, Ramasharma
TITLE OF INVENTION: METHODS AND
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94306-1546
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,847
FILING DATE: 27-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0009.31
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT
US-08-496-847-25

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 R 1

Search completed: February 25, 2004, 15:46:57
Job time : 4.79114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:44:50 ; Search time 5.8481 Seconds
(without alignments)
252.744 Million cell updates/sec

Title: US-09-877-606-8
Perfect score: 7
Sequence: 1 XXXXXXX 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------|--------------------|
| 1 | 0 | 0.0 | 1 | US-09-909-348-4 | Sequence 4, Appli |
| 2 | 0 | 0.0 | 1 | US-09-982-172-3 | Sequence 3, Appli |
| 3 | 0 | 0.0 | 1 | US-09-982-172-4 | Sequence 4, Appli |
| 4 | 0 | 0.0 | 1 | US-09-982-172-9 | Sequence 9, Appli |
| 5 | 0 | 0.0 | 1 | US-09-982-172-11 | Sequence 11, Appli |
| 6 | 0 | 0.0 | 1 | US-09-982-172-19 | Sequence 19, Appli |
| 7 | 0 | 0.0 | 1 | US-09-982-172-31 | Sequence 31, Appli |
| 8 | 0 | 0.0 | 1 | US-09-982-172-35 | Sequence 35, Appli |
| 9 | 0 | 0.0 | 1 | US-09-982-172-37 | Sequence 37, Appli |
| 10 | 0 | 0.0 | 1 | US-09-982-172-46 | Sequence 46, Appli |
| 11 | 0 | 0.0 | 1 | US-09-982-172-69 | Sequence 69, Appli |
| 12 | 0 | 0.0 | 1 | US-09-982-172-80 | Sequence 80, Appli |
| 13 | 0 | 0.0 | 1 | US-09-982-172-81 | Sequence 81, Appli |
| 14 | 0 | 0.0 | 1 | US-09-982-172-83 | Sequence 83, Appli |
| 15 | 0 | 0.0 | 1 | US-09-982-172-86 | Sequence 86, Appli |

| | | | | | |
|----|---|-----|---|--------------------|--------------------|
| 16 | 0 | 0.0 | 1 | US-09-982-172-93 | Sequence 93, Appli |
| 17 | 0 | 0.0 | 1 | US-09-982-172-95 | Sequence 95, Appli |
| 18 | 0 | 0.0 | 1 | US-09-982-172-106 | Sequence 106, App |
| 19 | 0 | 0.0 | 1 | US-09-982-172-112 | Sequence 112, App |
| 20 | 0 | 0.0 | 1 | US-09-982-172-120 | Sequence 120, App |
| 21 | 0 | 0.0 | 1 | US-09-982-172-126 | Sequence 126, App |
| 22 | 0 | 0.0 | 1 | US-09-982-172-148 | Sequence 148, App |
| 23 | 0 | 0.0 | 1 | US-09-982-172-149 | Sequence 149, App |
| 24 | 0 | 0.0 | 1 | US-09-982-172-155 | Sequence 155, App |
| 25 | 0 | 0.0 | 1 | US-09-982-172-160 | Sequence 160, App |
| 26 | 0 | 0.0 | 1 | US-09-982-172-172 | Sequence 172, App |
| 27 | 0 | 0.0 | 1 | US-09-982-172-173 | Sequence 173, App |
| 28 | 0 | 0.0 | 1 | US-09-982-172-175 | Sequence 175, App |
| 29 | 0 | 0.0 | 1 | US-09-982-172-189 | Sequence 189, App |
| 30 | 0 | 0.0 | 1 | US-09-982-172-190 | Sequence 190, App |
| 31 | 0 | 0.0 | 1 | US-09-982-172-191 | Sequence 191, App |
| 32 | 0 | 0.0 | 1 | US-09-982-172-195 | Sequence 195, App |
| 33 | 0 | 0.0 | 1 | US-09-982-172-200 | Sequence 200, App |
| 34 | 0 | 0.0 | 1 | US-09-982-172-211 | Sequence 211, App |
| 35 | 0 | 0.0 | 1 | US-09-809-391-395 | Sequence 395, App |
| 36 | 0 | 0.0 | 1 | US-09-809-391-611 | Sequence 611, App |
| 37 | 0 | 0.0 | 1 | US-09-882-171-395 | Sequence 395, App |
| 38 | 0 | 0.0 | 1 | US-09-882-171-611 | Sequence 611, App |
| 39 | 0 | 0.0 | 1 | US-09-833-245-184 | Sequence 184, App |
| 40 | 0 | 0.0 | 1 | US-09-833-245-186 | Sequence 186, App |
| 41 | 0 | 0.0 | 1 | US-09-833-245-325 | Sequence 325, App |
| 42 | 0 | 0.0 | 1 | US-09-833-245-500 | Sequence 500, App |
| 43 | 0 | 0.0 | 1 | US-09-833-245-744 | Sequence 744, App |
| 44 | 0 | 0.0 | 1 | US-09-833-245-1045 | Sequence 1045, Ap |
| 45 | 0 | 0.0 | 1 | US-09-833-245-1119 | Sequence 1119, Ap |

ALIGNMENTS

RESULT 1

US-09-909-348-4
; Sequence 4, Application US/09909348
; Patent No. US20020042373A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Stierberg, Janet
; APPLICANT: Bergmann, John
; TITLE OF INVENTION: Stimulation of Cartilage Growth With Agonists
; FILE REFERENCE: Of The No. US20020042373A1-Proteolytically Activated Thrombin R
; FILE REFERENCE: 3033.1003-001
; CURRENT APPLICATION NUMBER: US/09/909,348
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,800
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
; NAME/KEY: VARIANT
; LOCATION: (1)...(14)
; OTHER INFORMATION: Xaa at position six is Glu or Gln
; OTHER INFORMATION: Xaa at position thirteen is Phe, Met, Leu, His or
US-09-909-348-4

Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 V 1

```
RESULT 2
US-09-982-172-3
; Sequence 3, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-3
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 3
US-09-982-172-4
; Sequence 4, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-4
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 4
US-09-982-172-9
; Sequence 9, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
```

```
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-9
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 5
US-09-982-172-11
; Sequence 11, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERA
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-11
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 6
US-09-982-172-19
; Sequence 19, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERA
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-19
```

Query Match 0.0%; Score 0; DB 9; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 K 1

RESULT 7
 US-09-882-172-31
 ; Sequence 31, Application US/09982172
 ; Patent No. US20020137119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emil Israel Katz
 ; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
 ; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
 ; TITLE OF INVENTION: UTILIZING EACH
 ; FILE REFERENCE: 01/22283
 ; CURRENT APPLICATION NUMBER: US/09/982,172
 ; CURRENT FILING DATE: 2001-10-19
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 31
 ; LENGTH: 1
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Computer generated synthetic peptide
 US-09-882-172-31

Query Match 0.0%; Score 0; DB 9; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 K 1

RESULT 8
 US-09-882-172-35
 ; Sequence 35, Application US/09982172
 ; Patent No. US20020137119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emil Israel Katz
 ; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
 ; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
 ; TITLE OF INVENTION: UTILIZING EACH
 ; FILE REFERENCE: 01/22283
 ; CURRENT APPLICATION NUMBER: US/09/982,172
 ; CURRENT FILING DATE: 2001-10-19
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 35
 ; LENGTH: 1
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Computer generated synthetic peptide
 US-09-882-172-35

Query Match 0.0%; Score 0; DB 9; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 R 1

RESULT 9

US-09-882-172-37
 ; Sequence 37, Application US/09982172
 ; Patent No. US20020137119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emil Israel Katz
 ; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
 ; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
 ; TITLE OF INVENTION: UTILIZING EACH
 ; FILE REFERENCE: 01/22283
 ; CURRENT APPLICATION NUMBER: US/09/982,172
 ; CURRENT FILING DATE: 2001-10-19
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 37
 ; LENGTH: 1
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Computer generated synthetic peptide
 US-09-882-172-37

Query Match 0.0%; Score 0; DB 9; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 K 1

RESULT 10
 US-09-882-172-46
 ; Sequence 46, Application US/09982172
 ; Patent No. US20020137119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emil Israel Katz
 ; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
 ; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
 ; TITLE OF INVENTION: UTILIZING EACH
 ; FILE REFERENCE: 01/22283
 ; CURRENT APPLICATION NUMBER: US/09/982,172
 ; CURRENT FILING DATE: 2001-10-19
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 46
 ; LENGTH: 1
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Computer generated synthetic peptide
 US-09-882-172-46

Query Match 0.0%; Score 0; DB 9; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 K 1

RESULT 11
 US-09-882-172-69
 ; Sequence 69, Application US/09982172
 ; Patent No. US20020137119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emil Israel Katz
 ; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
 ; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
 ; TITLE OF INVENTION: UTILIZING EACH
 ; FILE REFERENCE: 01/22283
 ; CURRENT APPLICATION NUMBER: US/09/982,172
 ; CURRENT FILING DATE: 2001-10-19

; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-69

Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 K 1

RESULT 12
US-09-982-172-80
; Sequence 80, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-80

Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 K 1

RESULT 13
US-09-982-172-81
; Sequence 81, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-81

Query Match 0.0%; Score 0; DB 9; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 14
US-09-982-172-83
; Sequence 83, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-83

Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 15
US-09-982-172-86
; Sequence 86, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-86

Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 K 1

Search completed: February 25, 2004, 16:04:26
Job time : 5.8481 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: February 25, 2004, 15:36:14 ; Search time 2.25949 Seconds
(without alignments)
298.005 Million cell updates/sec

Title: US-09-877-606-8
Perfect score: 7
Sequence: 1 XXXXXXX 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 0 | 0.0 | 3 | 3 RHTDFO | thyroliberin - Bom |
| 2 | 0 | 0.0 | 3 | 3 RHPCT | thyroliberin - pig |
| 3 | 0 | 0.0 | 3 | 3 RSHST | thyroliberin - she |
| 4 | 0 | 0.0 | 3 | 3 A92971 | thyroliberin - eas |
| 5 | 0 | 0.0 | 3 | 3 GKHU | growth-modulating |
| 6 | 0 | 0.0 | 3 | 3 A60898 | bursin - chicken |
| 7 | 0 | 0.0 | 3 | 3 A23751 | spinal cord peptid |
| 8 | 0 | 0.0 | 3 | 3 B23751 | spinal cord peptid |
| 9 | 0 | 0.0 | 3 | 3 A33802 | thyrotropin-releas |
| 10 | 0 | 0.0 | 3 | 3 A22565 | R-phycoerythrin al |
| 11 | 0 | 0.0 | 3 | 3 PQ0010 | angiotensin-conver |
| 12 | 0 | 0.0 | 3 | 3 S13894 | histidinol dehydro |
| 13 | 0 | 0.0 | 3 | 3 A43391 | TRH-like tripeptid |
| 14 | 0 | 0.0 | 3 | 3 E37196 | bradykinin-potenti |
| 15 | 0 | 0.0 | 3 | 3 F37196 | bradykinin-potenti |
| 16 | 0 | 0.0 | 3 | 3 I50412 | gene p20K protein |
| 17 | 0 | 0.0 | 3 | 3 PT0636 | T-cell receptor be |
| 18 | 0 | 0.0 | 3 | 3 PT0578 | T-cell receptor be |
| 19 | 0 | 0.0 | 3 | 3 PT0571 | T-cell receptor be |
| 20 | 0 | 0.0 | 3 | 3 PT0622 | T-cell receptor be |
| 21 | 0 | 0.0 | 3 | 3 I78890 | tyrosine protein k |
| 22 | 0 | 0.0 | 3 | 3 S68328 | blood cell protein |
| 23 | 0 | 0.0 | 3 | 3 T13892 | cytochrome-c oxida |
| 24 | 0 | 0.0 | 4 | 1 ECXAA | antho-RFamide neur |
| 25 | 0 | 0.0 | 4 | 2 S18401 | thyroglobulin - do |
| 26 | 0 | 0.0 | 4 | 2 A02147 | phagocytosis-stimu |
| 27 | 0 | 0.0 | 4 | 2 A32039 | tyrosine-melanocyt |
| 28 | 0 | 0.0 | 4 | 2 ECNK | cardioexcitatory n |
| 29 | 0 | 0.0 | 4 | 2 PL0140 | carbon-monoxide de |

30 0 0.0 4 2 PL0146 carbon-monoxide de
31 0 0.0 4 2 A37832 phenol 2-monooxyge
32 0 0.0 4 2 A48360 gamma subunit of p
33 0 0.0 4 2 I40697 biotin A - Citroba
34 0 0.0 4 2 A61300 22K superhelical D
35 0 0.0 4 2 I57745 D-mannonate hydrol
36 0 0.0 4 2 A41890 protein D - Escher
37 0 0.0 4 2 S43014 hypothetical prote
38 0 0.0 4 2 D41654 hypothetical prote
39 0 0.0 4 2 B43848 cell surface adhes
40 0 0.0 4 2 I40505 hypothetical prote
41 0 0.0 4 2 I40870 phospholipase C (E
42 0 0.0 4 2 I40804 endoglucanase F -
43 0 0.0 4 2 T46627 starvation-induced
44 0 0.0 4 2 S53508 ribosomal protein
45 0 0.0 4 2 S17255

ALIGNMENTS

RESULT 1

RHTDFO
thyroliberin - Bombina orientalis
C;Species: Bombina orientalis
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A90919; A01415
R;Yasuhara, T.; Nakajima, T.
Chem. Pharm. Bull. 23, 3301-3303, 1975
A;Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.
A;Reference number: A90919; MUID:76138399; PMID:815011
A;Accession: A90919
A;Molecule type: protein
A;Residues: 1-3 <YAS>
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 2

RHPCT
thyroliberin - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A01415
R;Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.
Biochemistry 9, 1103-1106, 1970
A;Title: Structure of porcine thyrotropin releasing hormone.
A;Reference number: A90560; MUID:70136150; PMID:4984938
A;Accession: A01415
A;Molecule type: protein
A;Residues: 1-3 <NAI>
R;Boier, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.
Biochem. Biophys. Res. Commun. 37, 705-710, 1969
A;Title: The identity of chemical and hormonal properties of the thyrotropin releasing
A;Reference number: A90167; MUID:70039904; PMID:4982117
A;Contents: annotation

A;Note: biological activities and Rf values (in 17 chromatographic systems) of the synt:
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 3
RHSHT
thyroliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A93750; A01415
R:Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.
Org. Mass Spectrom. 5, 221-228, 1971
A:Title: The elucidation of the primary structure of the hypothalamic thyroïd stimulating
A:Reference number: A93750
A:Accession: A93750
A:Molecule type: protein
A:Residues: 1-3 <DES>
R:Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.
Nature 226, 321-325, 1970
A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.
A:Reference number: A93161; MUID:70163386; PMID:4985794
A:Contents: annotation
A:Note: physicochemical characteristics and biological activities of the natural and syn
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 4
A92971
thyroliberin - eastern newt (tentative sequence)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A92971; A01415
R:Grimm-Jorgensen, Y.; McKelvy, J.F.
J. Neurochem. 23, 471-478, 1974
A:Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) bra
A:Reference number: A92971; MUID:75035605; PMID:4214528
A:Accession: A92971
A:Molecule type: protein
A:Residues: 1-3 <GRI>
A:Note: a peptide with the chromatographic and electrophoretic characteristics of thyrol
stidine, or glutamic acid
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic
F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 5
GKHU
growth-modulating peptide - human

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 6
A60898
bursin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A60898
R:Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.
Science 231, 997-999, 1986
A:Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone of t
A:Reference number: A60898; MUID:86122316; PMID:3484938
A:Accession: A60898
A:Molecule type: protein
A:Residues: 1-3 <AUD>
C:Superfamily: unassigned animal peptides
C:Keywords: amidated carboxyl end; hormone
F3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 K 1

RESULT 7
A23751
spinal cord peptide SCP-4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A23751
R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K
Arch. Biochem. Biophys. 240, 178-183, 1985
A:Reference number: A23751; MUID:85250425; PMID:4015098
A:Accession: A23751
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <HSI>
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 8

C:Species: Homo sapiens (man)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A01421
R:Schlesinger, D.H.; Pickart, L.; Thaler, M.M.
Experientia 33, 324-325, 1977
A:Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.
A:Reference number: A01421; MUID:77162369; PMID:858356
A:Accession: A01421
A:Molecule type: protein
A:Residues: 1-3 <SCH>
A:Note: this serum tripeptide is found to stimulate growth of some cell types and to inh
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 G 1

RESULT 6
A60898
bursin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A60898
R:Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.
Science 231, 997-999, 1986
A:Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone of t
A:Reference number: A60898; MUID:86122316; PMID:3484938
A:Accession: A60898
A:Molecule type: protein
A:Residues: 1-3 <AUD>
C:Superfamily: unassigned animal peptides
C:Keywords: amidated carboxyl end; hormone
F3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 K 1

RESULT 7
A23751
spinal cord peptide SCP-4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A23751
R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K
Arch. Biochem. Biophys. 240, 178-183, 1985
A:Reference number: A23751; MUID:85250425; PMID:4015098
A:Accession: A23751
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <HSI>
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 8

B23751
spinal cord peptide SCP-5 - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: B23751
R;Hei, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.
Arch. Biochem. Biophys. 240, 178-183, 1985
A;Reference number: A23751; MUID:85250425; PMID:4015098
A;Accession: B23751
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3 <HS1>
C;Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 M 1

RESULT 9
A33802
thyrotropin-releasing hormone-like peptide - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A33802
R;Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.
J. Biol. Chem. 264, 7788-7791, 1989
A;Title: A novel peptide, pyroglutamylglutamylproline amide, in the rabbit prostate comp
A;Reference number: A33802; MUID:89255196; PMID:2498305
A;Accession: A33802
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3 <COC>
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 10
A22565
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3896644
A;Accession: A22565
A;Molecule type: protein
A;Residues: 1-3 <KIO>

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 2 Y 2

RESULT 11
PQ0010
angiotensin-converting enzyme inhibitor (FLP-3) - common fig
N;Alternate names: ficus latex peptide 3
C;Species: Ficus carica (common fig)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: PQ0010
R;Maruyama, S.; Miyoshi, S.; Tanaka, H.
Agric. Biol. Chem. 53, 2763-2767, 1989
A;Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.
A;Reference number: PQ0008
A;Accession: PQ0010
A;Molecule type: protein
A;Residues: 1-3 <MAR>
A;Experimental source: latex
C;Keywords: angiotensin-converting enzyme inhibitor

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 L 1

RESULT 12
S13894
histidinol dehydrogenase (EC 1.1.1.23) - wild cabbage (fragment)
C;Species: Brassica oleracea (wild cabbage)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: S13894
R;Nagai, A.; Scheidegger, A.
Arch. Biochem. Biophys. 284, 127-132, 1991
A;Title: Purification and characterization of histidinol dehydrogenase from cabbage.
A;Reference number: S13894; MUID:9112783; PMID:1989490
A;Accession: S13894
A;Molecule type: protein
A;Residues: 1-3 <NAG>
A;Experimental source: var. capitata
C;Keywords: dimer; NAD; oxidoreductase

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1
Db 1 X 1

RESULT 13
A43391
TRH-like tripeptide - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A43391
R;Jackey, D.B.
J. Biol. Chem. 267, 17508-17511, 1992
A;Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroglu
A;Reference number: A43391; MUID:92388092; PMID:1517203
A;Accession: A43391
A;Molecule type: protein
A;Residues: 1-3 <IAC>
C;Keywords: amidated carboxyl end; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 14

E37196
bradykinin-potentiating peptide 5 - island jararaca
C:Species: Bothrops insularis (island jararaca)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 30-Jun-2001
C:Accession: E37196
R:Cintrá, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides
A:Reference number: A37196; MUID:90351557; PMID:2386615
A:Accession: E37196
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <CIN>
C:Keywords: pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 15

F37196
bradykinin-potentiating peptide 6 - island jararaca
C:Species: Bothrops insularis (island jararaca)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: F37196
R:Cintrá, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides
A:Reference number: A37196; MUID:90351557; PMID:2386615
A:Accession: F37196
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <CIN>
C:Keywords: pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Q 1

Search completed: February 25, 2004, 15:45:42
Job time : 4.25949 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 15:32:39 ; Search time 1.41772 Seconds
(without alignments)
257.096 Million cell updates/sec

Title: US-09-877-606-8

Perfect score: 7
Sequence: 1 XXXXXXXX 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 0 | 0.0 | 3 | 1 GRWM_HUMAN | P01157 homo sapien |
| 2 | 0 | 0.0 | 3 | 1 LUXE_VIBFI | P24272 vibrio fisc |
| 3 | 0 | 0.0 | 3 | 1 THYL_PIG | P01151 sus scrofa |
| 4 | 0 | 0.0 | 4 | 1 ACHI_ACHFU | P35904 achatina fu |
| 5 | 0 | 0.0 | 4 | 1 DCML_PSECH | P19916 pseudomonas |
| 6 | 0 | 0.0 | 4 | 1 DCML_PSECH | P19918 pseudomonas |
| 7 | 0 | 0.0 | 4 | 1 EOSI_HUMAN | P02731 homo sapien |
| 8 | 0 | 0.0 | 4 | 1 FAR3_HIRME | P42562 hirudo medi |
| 9 | 0 | 0.0 | 4 | 1 FAR4_HIRME | P42563 hirudo medi |
| 10 | 0 | 0.0 | 4 | 1 FFKA_ANTEL | P58705 anthopleura |
| 11 | 0 | 0.0 | 4 | 1 FLRF_HIRME | P42561 hirudo medi |
| 12 | 0 | 0.0 | 4 | 1 FLRF_HIRME | P58707 anthopleura |
| 13 | 0 | 0.0 | 4 | 1 FMRE_MAGNI | P01162 macrocallis |
| 14 | 0 | 0.0 | 4 | 1 FYRI_ANTEL | P58706 anthopleura |
| 15 | 0 | 0.0 | 4 | 1 OCPI_OCTMI | P58648 octopus min |
| 16 | 0 | 0.0 | 4 | 1 OCP3_OCTMI | P58649 octopus min |
| 17 | 0 | 0.0 | 4 | 1 RM01_YEAST | P36515 saccharomyc |
| 18 | 0 | 0.0 | 4 | 1 TUFT_HUMAN | P01858 homo sapien |
| 19 | 0 | 0.0 | 5 | 1 AL14_CARMA | P81817 carcinus ma |
| 20 | 0 | 0.0 | 5 | 1 BIOA_CITFR | P13071 citrobacter |
| 21 | 0 | 0.0 | 5 | 1 BIOB_CITFR | P12997 citrobacter |
| 22 | 0 | 0.0 | 5 | 1 BPP7_BOTIN | P30425 bothrops in |
| 23 | 0 | 0.0 | 5 | 1 EI03_LITRU | P82099 litoria rub |
| 24 | 0 | 0.0 | 5 | 1 EI04_LITRU | P82100 litoria rub |
| 25 | 0 | 0.0 | 5 | 1 FARP_ARITR | P41853 artiposthi |
| 26 | 0 | 0.0 | 5 | 1 PAP2_PARMA | P81864 pardachirus |
| 27 | 0 | 0.0 | 5 | 1 PRCT_PERAM | P01373 periplaneta |
| 28 | 0 | 0.0 | 5 | 1 PSK_DAUCA | P58261 daucus caro |
| 29 | 0 | 0.0 | 5 | 1 RE11_LITRU | P82071 litoria rub |
| 30 | 0 | 0.0 | 5 | 1 RE21_LITRU | P82072 litoria rub |
| 31 | 0 | 0.0 | 5 | 1 RE31_LITRU | P82073 litoria rub |
| 32 | 0 | 0.0 | 5 | 1 RE32_LITRU | P82073 litoria rub |
| 33 | 0 | 0.0 | 5 | 1 SUGA_ACHDO | P19991 acheta dome |

ALIGNMENTS

RESULT 1

GRWM_HUMAN STANDARD; PRT; 3 AA.
AC P01157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Growth-modulating peptide.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77162369; PubMed=858356;
RA Schlesinger D.H., Pickart L., Thaler M.M.;
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";
RL Experientia 33:324-325(1977).
CC -|- MISCELLANEOUS: This serum tripeptide has been found to stimulate growth of some cell types and to inhibit other types in vitro.
CC GO; GO:0001558; P:regulation of cell growth; NAS.
SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

RESULT 2

LUXE_VIBFI STANDARD; PRT; 3 AA.
AC P24272;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-protein synthetase) (Fragment).
GN LUXE.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91072226; PubMed=2254256;
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon.";
RL J. Bacteriol. 172:6797-6802(1990).
CC -|- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID. IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS

34 0 0.0 5 1 TPIS_CANFA P54714 canis famil
35 0 0.0 5 1 TRM3_ECOLI P13973 escherichia
36 0 0.0 5 1 UC22_MAIZE P80628 zea mays (m
37 0 0.0 5 1 UF01_MOUSE P38639 mus musculu
38 0 0.0 5 1 UXA4_CHLTR P38005 chlamydia t
39 0 0.0 6 1 ACPH_RABIT P25154 oryctolagus
40 0 0.0 6 1 ASP2_LACSN P82655 lactobacill
41 0 0.0 6 1 CIP1_MYTED P13736 mytilus edu
42 0 0.0 6 1 CIP2_MYTED P13737 mytilus edu
43 0 0.0 6 1 EI01_LITRU P82096 litoria rub
44 0 0.0 6 1 FARP_MONEX P41966 moniezia ex
45 0 0.0 6 1 LOK1_LOCOMI P41491 locusta mig

CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
 CC -1- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
 CC an acyl-protein thioester.
 CC -1- PATHWAY: Bioluminescent fatty acid reduction system; second step.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M62812; -; NOT_ANNOTATED_CDS.
 CC Luminescence; Ligase.
 CC NON TER 1
 CC SEQUENCE 3 AA; 374 MW; 6AA33030000000000 CRC64;
 CC -----
 CC Query Match 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity 0.0%; Pred. No. 0;
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 X 1
 CC Db 1 I 1
 CC
 CC RESULT 3
 CC THYL_PIG STANDARD; PRT; 3 AA.
 CC AC P01151;
 CC DT 21-JUL-1986 (Rel. 01, Created)
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
 CC OS Sus scrofa (Pig),
 CC OS Ovis aries (Sheep),
 CC OS Bombina orientalis (Oriental fire-bellied toad), and
 CC OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC OC NCBI_TaxID=9823, 9940, 8346, 8316;
 CC [1]
 CC RN
 CC RP SEQUENCE.
 CC RC SPECIES=Pig; TISSUE=Hypothalamus;
 CC RX MEDLINE=70136150; PubMed=4984938;
 CC RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
 CC "Structure of porcine thyrotropin releasing hormone."
 CC RT Biochemistry 9:1103-1106(1970).
 CC RL [2]
 CC RN
 CC RP SYNTHESIS.
 CC RC SPECIES=Pig;
 CC RX MEDLINE=70039904; PubMed=4982117;
 CC RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
 CC "The identity of chemical and hormonal properties of the thyrotropin
 CC releasing hormone and pyroglutamyl-histidyl-proline amide."
 CC RT Biochem. Biophys. Res. Commun. 37:705-710(1969).
 CC RL [3]
 CC RN
 CC RP SEQUENCE.
 CC RC SPECIES=Sheep; TISSUE=Hypothalamus;
 CC RA Desiderio D.M. Jr., Burgus R., Dunn T.P., Vale W., Guillemin R.,
 CC Ward D.N.;
 CC "The elucidation of the primary structure of the hypothalamic thyroid
 CC stimulating hormone releasing factor of ovine origin by means of mass
 CC spectrometry."
 CC RT Org. Mass Spectrom. 5:221-228(1971).
 CC RL [4]
 CC RN
 CC RP SYNTHESIS.
 CC RC SPECIES=Sheep;
 CC RX MEDLINE=70163386; PubMed=4985794;
 CC RA Burgus R., Dunn T.P., Desiderio D.M., Ward D.N., Vale W.,
 CC Guillemin R.;
 CC "Characterization of ovine hypothalamic hypophysiotropic

RT TSH-releasing factor."
 RL Nature 226:321-325(1970).
 RN [5]
 RN SEQUENCE.
 RC SPECIES=B.orientalis; TISSUE=Skin;
 RX MEDLINE=76138399; PubMed=815011;
 RA Yasuhara T., Nakajima T.;
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin."
 RL Chem. Pharm. Bull. 23:3301-3303(1975).
 RN [6]
 RN SEQUENCE.
 RC SPECIES=N.viridescens;
 RX MEDLINE=75035605; PubMed=4214528;
 RA Grimm-Joergensen Y., McKelvy J.F.;
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
 RT viridescens) brain in vitro. Isolation and characterization of
 RT thyrotropin releasing factor."
 RL J. Neurochem. 23:471-478(1974).
 CC -1- FUNCTION: TRH functions as a regulator of the biosynthesis of TSH
 CC neuromodulator in the central and peripheral nervous systems.
 CC PIR; A90919; RHTOTO.
 CC DR PIR; A92971; A92971.
 CC KW Amidation; Pyrrolidone carboxylic acid.
 CC FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 3 3 AMIDATION.
 CC SQ SEQUENCE 3 AA; 380 MW; 7761F6B0000000000 CRC64;
 CC
 CC Query Match 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity 0.0%; Pred. No. 0;
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 X 1
 CC Db 1 Q 1
 CC
 CC RESULT 4
 CC ACHI_ACHFU STANDARD; PRT; 4 AA.
 CC AC P35904;
 CC DT 01-JUN-1994 (Rel. 29, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE Achatina-I.
 CC OS Achatina fulica (Giant African snail).
 CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
 CC OC NCBI_TaxID=6530;
 CC [1]
 CC RN
 CC RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 CC RC STRAIN=Perussac; TISSUE=Ganglion;
 CC RX MEDLINE=89273551; PubMed=2597281;
 RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
 RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
 RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
 RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 RT fulica Perussac containing a D-amino acid residue."
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 RN [2]
 RN CHARACTERIZATION.
 RC STRAIN=Perussac; TISSUE=Heart atrium;
 RX MEDLINE=91264856; PubMed=1675568;
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 RT "Purification of achatin-I from the atria of the African giant snail,
 RT Achatina fulica, and its possible function."
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 RN [3]
 RN X-RAY CRYSTALLOGRAPHY.
 RC MEDLINE=93014529; PubMed=1399265;
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

RA Iwashita T., Nomoto K.;
 RT "Crystal structure and molecular conformation of actinin-I
 RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
 RT D-amino acid residue.";
 RL Int. J. Pept. Protein Res. 39:258-264 (1992).
 CC -!- FUNCTION: Neuroexcitatory peptide; increases the impulse frequency
 CC and produces a spike broadening of the identified heart excitatory
 CC neuron (PON); also enhances the amplitude and frequency of the
 CC heart beat. Has also an effect on several other muscles.
 DR PIR; A32480; A32480.
 KW Hormone; D-amino acid.
 FT MOD RES 2 D-PHENYLALANINE.
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 3 A 3

RESULT 5

DCML_PSECH STANDARD; PRT; 4 AA.
 ID AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
 DE dehydrogenase subunit L) (CO-DH L) (Fragment).
 GN CUTL.
 OS Pseudomonas carboxydohydrogena.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae.
 OX NCBI_TaxID=290;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 RT carboxydohydrogenic bacteria.";
 RL Arch. Microbiol. 152:335-341 (1989).
 CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 CC dioxide.
 CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 CC acceptor.
 CC -!- COFACTOR: Molybdenum (molybdopterin).
 CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR; P19140; P19140.
 KW Oxidoreductase; Molybdenum.
 FT NON_TER 4
 SQ SEQUENCE 4 AA; 441 MW; 7761E876F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 M 1

RESULT 6

DCMS_PSECH STANDARD; PRT; 4 AA.
 ID AC P19918;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO

DE dehydrogenase subunit S) (CO-DH S) (Fragment).
 GN CUTS.
 OS Pseudomonas carboxydohydrogena.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae.
 OX NCBI_TaxID=290;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 RT carboxydohydrogenic bacteria.";
 RL Arch. Microbiol. 152:335-341 (1989).
 CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 CC dioxide.
 CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 CC acceptor.
 CC -!- COFACTOR: Binds 2 2Fe-2S clusters.
 CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR; P19146; P19146.
 KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
 FT NON_TER 4
 SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 2 A 2

RESULT 7

EOSI_HUMAN STANDARD; PRT; 4 AA.
 ID AC P02731;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Eosinophilic peptides.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=76078412; PubMed=1060093;
 RA Goetzl E.J., Austen K.F.;
 RT "Purification and synthesis of eosinophilic chemotactic factor of
 RT human lung tissue: identification as eosinophil chemotactic factor of
 RT anaphylaxis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127 (1975).
 CC -!- MISCELLANEOUS: These peptides are released from mast cells in lung
 CC (and other tissues) during hypersensitivity reactions
 CC (anaphylaxis). Their activities, preferentially affecting
 CC eosinophils, include chemotaxis, chemotactic deactivation, release
 CC of enzymes, and stimulation of the hexose monophosphate shunt.
 DR GO; GO:0006935; P:chemotaxis; IDA.
 DR GO; GO:0006955; P:immune response; IDA.
 FT VARIANT 1 V -> A (IN OTHER PEPTIDE).
 FT /FTID=VAR_005201.
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 3 S 3

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Antho-Kamide;
OS Anthopleura elegantissima (Sea anemone);
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaeae; Actinidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RN SEQUENCE
RX MEDLINE=92028852; PubMed=1681803;
RA Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenylactyl-Phe-Lys-Ala-NH2 (Antho-Kamide), a
RT novel neuropeptide from sea anemones.";
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN [2]
RN FUNCTION.
RX MEDLINE=93391436; PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Kamide and Antho-Riamide.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron specific.
DR PIR; JQ1273; JQ1273.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 512 MW; 6DD339C9A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 4 A 4

RESULT 11
FLRF_HIRME STANDARD; PRT; 4 AA.
AC P42561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FLRFamide.
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421, 27815;
RN [1]
RN SEQUENCE
RC SPECIES=H.medicalialis;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of RFamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [2]
RN SEQUENCE.
RC SPECIES=H.trivolvis; TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT trivolvis";
RL Peptides 15:31-36(1994).
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 4 AA.

RESULT 9
FAR4_HIRME STANDARD; PRT; 4 AA.
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide YMRP-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RN SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of RFamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 616 MW; 69D4068B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 Y 1

RESULT 10
FFKA_ATEL STANDARD; PRT; 4 AA.
AC P58705;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

```


Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 F 1

RESULT 12

FLRN ANTEL STANDARD; PRT; 4 AA.
 AC P58707; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Antho-RNamide.
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nyantheae; Actiniidae; Anthopleura.
 OX NCBI_TaxID=6110;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RX MEDLINE=90319122; PubMed=1973541;
 RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
 RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;
 RT "Isolation of L-3-phenylacetyl-Leu-Arg-Asp-NH2 (Antho-RNamide), a sea
 RT anemone neuropeptide containing an unusual amino-terminal blocking
 RT group.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414 (1990).
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- TISSUE SPECIFICITY: Neuron specific.
 CC -I- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.
 DR PIR; A35779; A35779.
 KW Neuropeptide; Amidation.
 FT MOD_RES 1 1 L-3-PHENYLACTYL.
 FT MOD_RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 F 1

RESULT 13

FMRF MACNI STANDARD; PRT; 4 AA.
 AC P01162;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE FMRFamide (Peak C) (Cardioexcitatory neuropeptide).
 OS Macrocallista nimbosa (Sun-ray clam),
 OS Nereis virens (Sandworm).
 OS Hirudo medicinalis (Medicinal leech), and
 OS Helisoma trivolvis (Snail).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
 OC Veneroidea; Veneridae; Macrocallista.
 OX NCBI_TaxID=6594, 6353, 6421, 27815;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;
 RX MEDLINE=77215956; PubMed=877582;
 RA Price D.A., Greenberg M.J.;
 RT "Structure of a molluscan cardioexcitatory neuropeptide.";
 RL Science 197:670-671 (1977).
 RN [2]
 RP SEQUENCE, AND CHARACTERIZATION.

RC SPECIES=M.nimbosa; TISSUE=Ganglion;
 RX MEDLINE=78012038; PubMed=909875;
 RA Price D.A., Greenberg M.J.;
 RT "Purification and characterization of a cardioexcitatory neuropeptide
 RT from the central ganglia of a bivalve mollusc.";
 RL Prep. Biochem. 7:261-281 (1977).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=N.virens;
 RX MEDLINE=90259866; PubMed=2342992;
 RA Krajniak K.G., Price D.A.;
 RT "Authentic FMRFamide is present in the polychaete Nereis virens.";
 RL Peptides 11:75-77 (1990).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=H.medicinalis;
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of RFamide neuropeptides in the medicinal leech.";
 RL Peptides 12:897-908 (1991).
 RN [5]
 RP SEQUENCE.
 RC SPECIES=H.trivolvis; TISSUE=Kidney;
 RX MEDLINE=94286417; PubMed=7912428;
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
 RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
 RT trivolvis.";
 RL Peptides 15:31-36 (1994).
 CC -I- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
 CC activities include augmentation, induction, and regularization of
 CC cardiac contraction.
 CC -I- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
 CC family.
 DR PIR; A01426; ECKN.
 DR PIR; A60418; A60418.
 KW Neuropeptide; Amidation.
 FT MOD_RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 Db 1 F 1

RESULT 14

FYRI ANTEL STANDARD; PRT; 4 AA.
 AC P58706;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Antho-Riamide I [Contains: Antho-Riamide II].
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nyantheae; Actiniidae; Anthopleura.
 OX NCBI_TaxID=6110;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92270459; PubMed=1821096;
 RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
 RA Grimmelikhuijzen C.J.P.;
 RT "Isolation of two novel neuropeptides from sea anemones: the unusual,
 RT biologically active L-3-phenylacetyl-Tyr-Arg-Ile-NH2 and its
 RT des-phenylacetyl fragment Tyr-Arg-Ile-NH2.";
 RL Peptides 12:1165-1173 (1991).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=93391436; PubMed=8397415;
 RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;

RT "The expansion behaviour of sea anemones may be coordinated by two
 RT inhibitory neuropeptides, Antho-Kaamide and Antho-Riamide.";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
 CC -I- FUNCTION: Inhibits spontaneous contractions in several muscle
 CC groups. May be involved in the expansion phase of feeding
 CC behaviour in sea anemones.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- TISSUE SPECIFICITY: Neuron specific.
 KW Neuropeptide; Amidation.
 FT CHAIN 1 4 ANTHO-RIAMIDE I.
 FT CHAIN 2 4 ANTHO-RIAMIDE II.
 FT MOD_RES 1 1 L-3-PHENYLACTYL.
 FT MOD_RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred.No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 F 1

RESULT 15

OCPI_OCTMI STANDARD; PRT; 4 AA.
 AC P58648;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Cardioactive peptides Ocp-1/Ocp-2.
 OS Octopus minor (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=89766;
 RN [1]
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hisada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 RT Octopus minor";
 RL Peptides 21:623-630(2000).
 CC -I- FUNCTION: Cardioactive; has both positive chronotropic and
 CC inotropic effects on the heart. Ocp-2 is a 1000 time less
 CC active than Ocp-1.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- PTM: Ocp-2 has L-Phe instead of D-Phe.
 CC -I- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
 KW Hormone; D-amino acid.
 FT MOD_RES 2 2 D-PHENYLALANINE (IN OCP-1).
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred.No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 G 1

Search completed: February 25, 2004, 15:41:51
 Job time : 4.41772 secs

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|--------|----|--------|-------------|
| | | Match | Length | | | |
| 1 | 0 | 0.0 | 2 | 5 | P83570 | sepia offic |
| 2 | 0 | 0.0 | 4 | 5 | P83568 | sepia offic |
| 3 | 0 | 0.0 | 4 | 11 | Q08433 | rattus sp. |
| 4 | 0 | 0.0 | 2 | 2 | P83073 | bacillus ce |
| 5 | 0 | 0.0 | 5 | 10 | Q99007 | hordeum vu |
| 6 | 0 | 0.0 | 5 | 13 | P83308 | gallus gall |
| 7 | 0 | 0.0 | 6 | 2 | P83533 | lactobacill |
| 8 | 0 | 0.0 | 6 | 5 | P83569 | sepia offic |
| 9 | 0 | 0.0 | 6 | 10 | P82181 | spinacia o |
| 10 | 0 | 0.0 | 6 | 10 | P82541 | spinacia o |
| 11 | 0 | 0.0 | 6 | 10 | P82182 | spinacia o |
| 12 | 0 | 0.0 | 7 | 2 | Q8KMS3 | klebsiella |
| 13 | 0 | 0.0 | 7 | 2 | Q47505 | escherichia |
| 14 | 0 | 0.0 | 7 | 2 | P70804 | azotobacter |
| 15 | 0 | 0.0 | 7 | 2 | O50556 | actinobacil |
| 16 | 0 | 0.0 | 7 | 2 | Q47477 | escherichia |

RESULT 2
 P83568 ID P83568 PRELIMINARY; PRT; 4 AA.
 AC P83568;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Pheromone peptide ILME.
 OS Sepia officinalis (Common cuttlefish).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Decapodiformes; Sepioidae; Sepiidae; Sepia.
 OX NCBI_TaxID=6610;
 RN [1]
 RP SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MASS
 RP SPECTROMETRY.
 RC TISSUE=Egg;
 RX PubMed=10944467;
 RA Zatylny C., Gagnon J., Boucaud-Camou E., Henry J.;
 RT "ILME: a waterborne pheromonal peptide released by the eggs of Sepia
 RT officinalis.";
 RL Biochem. Biophys. Res. Commun. 275:217-222(2000).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Egg;
 RX PubMed=12207899;
 RA Zatylny C., Marvin L., Gagnon J., Henry J.;
 RT "Fertilization in Sepia officinalis: the first mollusk sperm-
 RT attracting peptide.";
 RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).
 CC 1- FUNCTION: HAS MYOTROPIC ACTIVITY TARGETING THE GENITAL TRACT.
 CC 1- SUBCELLULAR LOCATION: SECRETED.
 CC 1- TISSUE SPECIFICITY: FOLLICULAR, FULLY GROWN OOCYTE AND EGG(EC2).
 CC 1- MASS SPECTROMETRY: MW=505.4; METHOD=WALDI.
 DR GO; GO:0005186; F:pheromone activity; IEA.
 KW Pheromone.
 SQ SEQUENCE 4 AA; 505 MW; 6B16972030000000 CRC64;

 Query Match 0.0%; Score 0; DB 5; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 X 1
 DB 1 I 1

 RESULT 3
 Q08433 ID Q08433 PRELIMINARY; PRT; 4 AA.
 AC Q08433;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Bilirubin UDP-glucuronosyltransferase (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Gunn;
 RX MEDLINE=91282758; PubMed=1840486;
 RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
 RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
 RT hyperbilirubinemic Gunn rat.";
 RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
 DR EMBL; S38636; AAB19259.1; -;
 DR GO; GO:0016740; F:transferase activity; IEA.
 KW Transferase.
 FT NON TER 1
 SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

 Query Match 0.0%; Score 0; DB 11; Length 4;

Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 X 1
 DB 1 N 1

 RESULT 4
 P83073 ID P83073 PRELIMINARY; PRT; 5 AA.
 AC P83073;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 88 kDa protein (Fragment).
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NCIMB 11796;
 RA Browne N., Dowds B.C.A.;
 RL Submitted (JUL-2001) to Swiss-Prot.
 FT NON TER 5
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;

 Query Match 0.0%; Score 0; DB 2; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 X 1
 DB 4 T 4

 RESULT 5
 Q99007 ID Q99007 PRELIMINARY; PRT; 5 AA.
 AC Q99007;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Alpha amylase (Fragment).
 GN AMY1 GENE.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91329704; PubMed=1831055;
 RA Jacobsen J.V., Close T.J.;
 RT "Control of transient expression of chimaeric genes by gibberellic
 RT acid and abscisic acid in protoplasts prepared from mature barley
 RL Plant Mol. Biol. 16:713-721(1991).
 DR EMBL; X54643; CAA38455.1; -;
 FT NON TER 5
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;

 Query Match 0.0%; Score 0; DB 10; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 X 1
 DB 2 A 2

 RESULT 6
 P83308

ID P83308 PRELIMINARY; PRT; 5 AA.
AC P83308; 2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DE FMRamide-like neuropeptide (LPLRF-amide).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Brain;
RX PubMed=6137771;
RA Dockray G.J.; Reeve J.R. Jr.; Shively J.; Gayton R.J.; Barnard C.S.;
RT "A novel active pentapeptide from chicken brain identified by
RT antibodies to FMRamide";
RL Nature 305:328-330(1983).
CC -1- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
CC FAMILY.
CC GO: GO:0007218; P:neuropeptide signaling pathway; TAS.
KW Neuropeptide; Amidation.
FT MOD_RES 5 5
SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;
Query Match 0.0%; Score 0; DB 13; Length 5;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 L 1

RESULT 7
ID P83533 PRELIMINARY; PRT; 6 AA.
AC P83533;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DE Unknown protein from 2D-page (Fragment).
OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1625;
RN [1]
RP SEQUENCE.
RC STRAIN=DSM 20451;
RX PubMed=12112860;
RA Drews O.; Weiss W.; Reil G.; Parlar H.; Wait R.; Goerg A.;
RT "High pressure effects step-wise altered protein expression in
RT Lactobacillus sanfranciscensis";
RL Proteomics 2:765-774(2002).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC PROTEIN IS: 15 KDA.
FT NON_TER 1 1
FT NON_TER 6 6
SQ SEQUENCE 6 AA; 590 MW; 6DDDD452D1AAC000 CRC64;
Query Match 0.0%; Score 0; DB 2; Length 6;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 2 T 2

RESULT 8
ID P83569 PRELIMINARY; PRT; 6 AA.
AC P83569;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DE Sperm attracting peptide SepSAF.
OS Sepia officinalis (Common cuttlefish).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Decapodiformes; Sepioidea; Sepiidae; Sepia.
OX NCBI_TaxID=6610;
RN [1]
RP SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND
RP AMIDATION.
RC TISSUE=Egg;
RX PubMed=12207899;
RA Zatylny C.; Marvin L.; Gagnon J.; Henry J.;
RT "Fertilization in Sepia officinalis: the first mollusk sperm-
RT attracting peptide";
RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).
CC -1- FUNCTION: ATTRACTS SPERM INCREASING THE CHANCES OF GAMETE
CC COLLISION.
CC -1- DEVELOPMENTAL STAGE: FIRST APPEARS IN THE OVARIAN FOLLICLES FURING
CC VITELLOGENESIS. ACCUMULATES IN THE OOCYTES BEFORE BEING SECRETED
CC DURING FERTILIZATION. EXPRESSION CONTINUES IN THE EMBEDDED OOCYTE.
CC ACCUMULATES IN THE EGG CAPSULE AFTER FERTILIZATION.
CC -1- MASS SPECTROMETRY: MW=596.6; METHOD=MALDI.
KW Amidation.
FT MOD_RES 6 6
SQ SEQUENCE 6 AA; 597 MW; 72C8676AA0470000 CRC64;
Query Match 0.0%; Score 0; DB 5; Length 6;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 2 I 2

RESULT 9
ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TREMELrel. 14, Created)
DT 01-JUN-2000 (TREMELrel. 14, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K.; Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR GO: GO:0009507; C:chloroplast; IEA.
DR GO: GO:0019843; F:rRNA binding; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro; IPR02363; Ribosomal L10eub.
DR PROSITE; PS01109; RIBOSOMAL L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

```

Query Match          0.0%; Score 0; DB 10; Length 6;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 A 1

RESULT 10
P82541
ID P82541 PRELIMINARY; PRT; 6 AA.
AC P82541;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS Spinacia oleracea (Spinach).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]

SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC STRAIN=cv. ALWARO; TISSUE=Leaf;
RX MEDLINE=20435797; PubMed=10874039;
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
the small subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 37:28455-28465 (2000).
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -|- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC -|- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC -|- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
FORM IS THE MINOR BASIC FORM.
CC -|- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
CC -|- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0019843; F:rRNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro; IPR002222; Ribosomal S19.
DR PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 732 MW; 6333735A41C000 CRC64;

Query Match          0.0%; Score 0; DB 10; Length 6;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 T 1

RESULT 11
P82182
ID P82182 PRELIMINARY; PRT; 6 AA.
AC P82182;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]

```

```

RP SEQUENCE.
RC STRAIN=cv. ALWARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482 (2000).
CC -|- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -|- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -|- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0019843; F:rRNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro; IPR002363; Ribosomal L10eub.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match          0.0%; Score 0; DB 10; Length 6;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 A 1

RESULT 12
Q8KMS3
ID Q8KMS3 PRELIMINARY; PRT; 7 AA.
AC Q8KMS3;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Putative MerR2 protein.
GN MERR2.
OS Klebsiella sp. LS13-39.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=143776;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS13-39;
RX MEDLINE=21604134; PubMed=11763242;
RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yurieva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
bacteria and their classification.";
RL Res. Microbiol. 152:811-822 (2001).
DR EMBL; AJ302776; CAC82975.1; -.
SQ SEQUENCE 7 AA; 608 MW; 6DC1B5BDD87DD6F0 CRC64;

Query Match          0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 2 A 2

RESULT 13
Q47505
ID Q47505 PRELIMINARY; PRT; 7 AA.
AC Q47505;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE McaA protein.

```

GN MCCA.
OS Escherichia coli.
OG Plasmid pMcC7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96099297; PubMed=8522520;
RA Gonzalez-Pastor J.R., San Millan J.L., Castilla M.A., Moreno F.;
RT "Structure and organization of plasmid genes required to produce the
RT translation inhibitor microcin C7";
RL J. Bacteriol. 177:7131-7140(1995).
DR EMBL; X57583; CAA40808.1; --
DR PIR; S45311; S45311.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 X 1

DB 3 T 3

RESULT 14

P70804
ID P70804 PRELIMINARY; PRT; 7 AA.
AC P70804;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Algt protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8830682;
RA Rehm B.H.A., Estesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; --
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 X 1

DB 2 T 2

RESULT 15

O50556
ID O50556 PRELIMINARY; PRT; 7 AA.
AC O50556;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GlyA (Fragment).
GN GLYA.
OS Actinobacillus actinomycetemcomitans (Haemophilus

